

to appear in the American Journal of Mathematics

GRADIENT ESTIMATES VIA NON-LINEAR POTENTIALS

FRANK DUZAAR AND GIUSEPPE MINGIONE

ABSTRACT. We present pointwise gradient bounds for solutions to p -Laplacean type non-homogeneous equations employing non-linear Wolff type potentials, and then prove similar bounds, via suitable caloric potentials, for solutions to parabolic equations.

CONTENTS

1. Introduction and main results	1
2. Notations	7
3. Elliptic estimates	9
4. Parabolic estimates	22
5. General weak, and very weak, solutions	32
6. Integrability estimates	35
References	38

1. INTRODUCTION AND MAIN RESULTS

The aim of this paper is to give a somehow unexpected but nevertheless natural maximal order - and parabolic - version of a by now classical result due to Kilpeläinen & Malý [35], later extended, by mean of a different approach, by Trudinger & Wang [56]. In [35, 56] the authors obtained a neat pointwise bound - see (1.14) below - for solutions to non-homogeneous quasi-linear equations of p -Laplacean type by mean of certain natural Wolff type non-linear potentials of the right-hand side measure datum; this is commonly considered as a basic result in the theory of quasi-linear equation. In this paper *we upgrade such a result by showing that similar pointwise bounds also hold for the gradient of solutions*, see (1.6) below. The resulting estimate finally allows to draw a complete and unified picture concerning gradient estimates for degenerate problems: it permits to recast classical gradient estimates for the non-homogeneous p -Laplacean equation obtained by Iwaniec [29] and DiBenedetto & Manfredi [18], as well as the pointwise L^∞ estimates obtainable when the right hand side is integrable enough [45, 46], up to the classical gradient bounds of Boccardo & Gallouët [10, 11], Lindqvist [42] and Dolzmann & Hungerbühler & Müller [19] for measure data problems; for more on the p -Laplacean see also the notes of Lindqvist [43]. More significantly, the results here allow for a *characterization of Lipschitz continuity of solutions in terms of Riesz potentials which is analogous to the one available for the standard Poisson equation*; see (1.8) and (1.16) below. Furthermore, the estimates extend to the case of coefficient dependent equations, allowing to get regularity criteria with respect to the regularity of the coefficients which are exactly those ones suggested by the analysis of the fundamental solutions of linear equations [25]. Our elliptic results find a natural but non-trivial analogue in the parabolic case, as we are also going to show here for a significant class of parabolic problems. In this case we shall employ a natural family of “caloric” type Riesz potentials to give a pointwise bound for the spatial derivative of solutions; in turn, the resulting inequality implies relevant

sharp borderline integral estimates which, known in the elliptic case, are otherwise unreachable via the known techniques, in the parabolic one. In particular, this finally allows to make the missing link between elliptic and parabolic borderline estimates such as those in Lorentz spaces. Moreover, as in the elliptic case, *we find a sharp characterization of the local boundedness of the gradient via Caloric potentials*; see (1.27) below. It is worth mentioning that the Lipschitz continuity estimates and the techniques of this paper are the starting point for further developments, eventually leading to C^1 -regularity assertions for solutions [23]. For notation and more details we recommend the reader to look at Section 2 below.

1.1. Degenerate Elliptic estimates. In this section the growth exponent p will be a number such that $p \geq 2$, while for the rest of the paper Ω will denote a bounded open and Lipschitz domain of \mathbb{R}^n , with $n \geq 2$. We shall consider general non-linear, possibly degenerate elliptic equations with p -growth of the type

$$(1.1) \quad -\operatorname{div} a(x, Du) = \mu.$$

whenever μ is a Radon measure defined on Ω with finite total mass; eventually letting $\mu(\mathbb{R}^n \setminus \Omega) = 0$, without loss of generality we may assume that μ is defined on the whole \mathbb{R}^n . The continuous vector field $a: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is assumed to be C^1 -regular in the gradient variable z , with $a_z(\cdot)$ being Carathéodory regular, and satisfying the following *growth, ellipticity and continuity assumptions*:

$$(1.2) \quad \begin{cases} |a(x, z)| + |a_z(x, z)|(|z|^2 + s^2)^{\frac{1}{2}} \leq L(|z|^2 + s^2)^{\frac{p-1}{2}} \\ \nu^{-1}(|z|^2 + s^2)^{\frac{p-2}{2}}|\lambda|^2 \leq \langle a_z(x, z)\lambda, \lambda \rangle \\ |a(x, z) - a(x_0, z)| \leq L_1\omega(|x - x_0|)(|z|^2 + s^2)^{\frac{p-1}{2}} \end{cases}$$

whenever $x, x_0 \in \Omega$ and $z, \lambda \in \mathbb{R}^n$, where $0 < \nu \leq L$ and $s \geq 0, L_1 \geq 1$ are fixed parameters. In (1.2) the function $\omega: [0, \infty) \rightarrow [0, \infty)$ is a modulus of continuity i.e., a non-decreasing concave function such that $\omega(0) = 0 = \lim_{\varrho \downarrow 0} \omega(\varrho)$ and $\omega(\cdot) \leq 1$. On such a function we impose a natural decay property, which is essentially optimal for the result we are going to have, and prescribes a *Dini-continuous dependence of the partial map* $x \mapsto a(x, z)/(|z| + s)^{p-1}$:

$$(1.3) \quad \int_0^R [\omega(\varrho)]^{\frac{2}{p}} \frac{d\varrho}{\varrho} := d(R) < \infty.$$

The prototype of (1.1) is - choosing $s = 0$ and omitting the x -dependence - clearly given by the p -Laplacean equation

$$(1.4) \quad -\operatorname{div} (|Du|^{p-2} Du) = \mu.$$

Our estimate involves the classical non-linear Wolff potential defined by

$$(1.5) \quad \mathbf{W}_{\beta, p}^\mu(x_0, R) := \int_0^R \left(\frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-\beta p}} \right)^{\frac{1}{p-1}} \frac{d\varrho}{\varrho} \quad \beta \in (0, n/p].$$

The first result, which we naturally state as an a priori estimate, is

Theorem 1.1 (Non-linear potential gradient bound). *Let $u \in C^1(\Omega)$ be a weak solution to (1.1) with $\mu \in L^1(\Omega)$ under the assumptions (1.2). Then there exists a constant $c \equiv c(n, p, \nu, L) > 0$, and a positive radius $\tilde{R} \equiv \tilde{R}(n, p, \nu, L, L_1, \omega(\cdot))$, such that the pointwise estimate*

$$(1.6) \quad |Du(x_0)| \leq c \int_{B(x_0, R)} (|Du| + s) dx + c \mathbf{W}_{\frac{1}{p}, p}^\mu(x_0, 2R)$$

holds whenever $B(x_0, 2R) \subseteq \Omega$ and $R \leq \tilde{R}$. Moreover, when the vector field $a(\cdot)$ is independent of x - and in particular for the p -Laplacean operator (1.4) - estimate (1.6) holds without any restriction on R .

The potential $\mathbf{W}_{\frac{1}{p},p}^\mu$ appearing in (1.6) is the natural one since its shape respects the scaling properties of the equation with respect to the estimate in question - see Section 1.2 and compare with the linear estimate in (1.12); moreover, the potential $\mathbf{W}_{\frac{1}{p},p}^\mu$ is in a sense *optimal*, as in an estimate like (1.6) it implies the recovering of the sharp integrability results known for general weak solutions to (1.1). Moreover, as a matter of fact when the right hand side is a properly singular measure estimate (1.6) reverses; for all such aspects we refer to Remark 6.2 below. An approximation procedure eventually allows to remove the extra regularity assumptions $u \in C^1(\Omega)$ and $\mu \in L^1(\Omega)$ up to consider the most general case of solutions to measure data problems as

$$(1.7) \quad \begin{cases} -\operatorname{div} a(x, Du) = \mu & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Our result indeed holds for general Solutions Obtained by Limit of Approximations (SOLA) to (1.7), a class of very weak solutions - solutions not necessarily lying in the natural energy space $W^{1,p}(\Omega)$ - which is introduced to get existence and uniqueness in several cases. Moreover, SOLA coincide with usual weak solutions for regular data $\mu \in W^{-1,p'}(\Omega)$; we refer to Section 5 below for more details, see also Theorems 5.1-5.2 below. When extended to general weak solutions *estimate (1.6) tells us the remarkable fact that the boundedness of Du at a point x_0 is independent of the solution u , and of the vector field $a(\cdot)$, but only depends on the behavior of $|\mu|$ in a neighborhood of x_0 .*

Corollary 1.1 ($C^{0,1}$ -regularity criterium). *Let $u \in W_0^{1,p-1}(\Omega)$ be a SOLA to the problem (1.7) - which is unique in the case $\mu \in L^1(\Omega)$ - under the assumptions (1.2) and (1.3). Then*

$$(1.8) \quad \mathbf{W}_{\frac{1}{p},p}^\mu(\cdot, R) \in L^\infty(\Omega) \text{ for some } R > 0 \implies Du \in L_{\text{loc}}^\infty(\Omega, \mathbb{R}^n).$$

Furthermore, there exists a constant c , depending only on $n, \nu, L, L_1, d(\cdot)$, such that the following estimate holds whenever $B_{2R} \subseteq \Omega$:

$$(1.9) \quad \|Du\|_{L^\infty(B_{R/2})} \leq c \int_{B(x_0, R)} (|Du| + s) dx + c \left\| \mathbf{W}_{\frac{1}{p},p}^\mu(\cdot, R) \right\|_{L^\infty(B_R)}.$$

Remark 1.1. Corollary 1.1 allows for a natural Lipschitz continuity criterium with respect to the regularity of coefficients (1.3) and moreover to obtain for the p -Laplacean, essentially the same criterium of Lipschitz continuity available for the Laplacean, that is the one via Riesz potentials; see (1.16) below and also (6.4) for a corollary. *We remark that finding conditions on μ implying the boundedness of Du was a major open issue in the theory of p -Laplacean equation and (1.8) seems to provide a first satisfying answer to the problem.* Indeed Cianchi [13] has shown that already in the plain case of the standard Poisson equation

$$(1.10) \quad -\Delta u = \mu$$

the gradient might be unbounded as soon as $\mathbf{W}_{\frac{1}{2},2} \notin L^\infty$. The estimates above also emphasize the role played by Dini-continuity of the coefficients in order to get the boundedness of the gradient, which is basically the same one observed when dealing with linear elliptic equations with variable Dini-continuous coefficients, and proving pointwise bounds on the derivatives of the Green's function [25, Section 3]. As a matter of fact a feature of the approach adopted here is that estimate (1.6) allows to derive conditions for gradient boundedness, *which are borderline simultaneously* w.r.t. the right hand side - i.e. (1.8) - and w.r.t. coefficients - i.e. (1.3), already in the

linear case. Indeed, a recent striking example of Jin & Mazya & Van Schaftingen [30] shows that there exist distributional solutions u to linear equations of the type

$$\operatorname{div}(A(x)Du) = 0$$

with a continuous (not Dini-continuous) and elliptic matrix $A(\cdot)$, such that $u \in W^{1,p}$ for every $p < \infty$ as predicted by the standard regularity theory, but such that $u \notin W^{1,\infty}$; in this case one even has that $Du \notin \operatorname{BMO}$.

1.2. Zero-order and related estimates. The linkage between estimate (1.6) and the estimates of [35, 56] is clear. The linear case $a(x, z) = z$ obviously leads to the standard Poisson equation (1.10) - here for simplicity considered in the whole \mathbb{R}^n - for which, due to the use of classical representation formulas, it is possible to get pointwise bounds for solutions via the use of Riesz potentials

$$(1.11) \quad I_\beta(\mu)(x) := \int_{\mathbb{R}^n} \frac{d\mu(y)}{|x-y|^{n-\beta}}, \quad \beta \in (0, n]$$

such as

$$(1.12) \quad |u(x)| \leq cI_2(|\mu|)(x) \quad \text{and} \quad |Du(x)| \leq cI_1(|\mu|)(x),$$

the first being actually valid for $n \geq 3$. We recall that the equivalent, localized version of the Riesz potential $I_\beta(\mu)$ is given by the linear potential

$$(1.13) \quad \mathbf{I}_\beta^\mu(x_0, R) \equiv \mathbf{W}_{\frac{\beta}{2}, 2}^\mu(x_0, R) = \int_0^R \frac{\mu(B(x_0, \varrho))}{\varrho^{n-\beta}} \frac{d\varrho}{\varrho} \leq I_\beta(\mu)(x_0),$$

the last line being valid for non-negative measures (note that we use two different - but similar - notations for the Riesz potential in (1.11) and its “polar version” in (1.13)). The natural non-linear version of (1.12), derived in [35, 56] for non-negative measures, when $p \leq n$, is treated via Wolff potentials:

$$(1.14) \quad |u(x_0)| \leq c \left(\int_{B(x_0, R)} (|u| + s)^\gamma dx \right)^{\frac{1}{\gamma}} + c \mathbf{W}_{1,p}^\mu(x_0, 2R) \quad \gamma > p - 1.$$

Estimate (1.6) upgrades (1.14) to the gradient/maximal level, obviously replacing $\mathbf{W}_{1,p}^\mu$ with $\mathbf{W}_{\frac{1}{p}, p}^\mu$, and represents the p -Laplacian analog of the second estimate in the left hand side of (1.12). Indeed, for $p = 2$ by (1.13) we have $\mathbf{W}_{\frac{1}{2}, 2}^\mu(x_0, R) \leq I_1(|\mu|)(x_0)$. Let us mention that the technique developed for estimate (1.6) also yields an alternative proof of estimate (1.14) which now holds for general signed measures; see Remark 4.1 below. We also refer to [49] for gradient potential estimates when $p = 2$, and to the important work of Labutin [38] for relevant Wolff type potential estimates related to fully non-linear Hessian type operators. For further interesting relations between degenerate quasilinear equations and potentials we refer to the recent interesting paper of Lindqvist & Manfredi [44]. Another consequence of estimate (1.6) and of the classical bound

$$(1.15) \quad \begin{aligned} \mathbf{W}_{\frac{1}{p}, p}^\mu(x_0, \infty) &= \int_0^\infty \left(\frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-1}} \right)^{\frac{1}{p-1}} \frac{d\varrho}{\varrho} \\ &\leq cI_{\frac{1}{p}} \left\{ \left[I_{\frac{1}{p}}(|\mu|) \right]^{\frac{1}{p-1}} \right\} (x_0) =: c\mathbf{V}_{\frac{1}{p}, p}(|\mu|)(x_0) \end{aligned}$$

is the estimate

$$(1.16) \quad |Du(x_0)| \leq c \int_{B(x_0, R)} (|Du| + s) dx + c\mathbf{V}_{\frac{1}{p}, p}(|\mu|)(x_0)$$

which holds whenever $B(x_0, 2R) \subseteq \Omega$ satisfies the smallness condition imposed in Theorem 1.1. Here we remind the reader that we have previously extended μ

to the whole space \mathbb{R}^n . The non-linear potential $\mathbf{V}_{\frac{1}{p},p}(\mu)(x_0)$ - often called the Havin-Maz'ja potential of μ - is a classical object in non-linear potential theory, and together with the bound (1.15) comes from the fundamental and pioneering work of Adams & Meyers and Havin & Maz'ja; see also [4, 5, 27]. Estimate (1.16) allows to derive all types of local estimates starting by the properties of the Riesz potential; see Section 6 below.

1.3. Parabolic estimates. Our aim here is not only to give a parabolic version of the elliptic estimate (1.6), but also to give a zero order estimate, that is the parabolic analog of the zero order elliptic estimate [35], which at this point will essentially follow as a corollary of the proof of the gradient estimate. We consider quasi-linear parabolic equations of the type

$$(1.17) \quad u_t - \operatorname{div} a(x, t, Du) = \mu,$$

in a cylindrical domain $\Omega_T := \Omega \times (-T, 0)$, where as in the previous sections $\Omega \subseteq \mathbb{R}^n$, $n \geq 2$ and $T > 0$. The vector-field $a: \Omega_T \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is assumed to be Carathéodory regular together with $a_z(\cdot)$, and indeed being C^1 -regular with respect to the gradient variable $z \in \mathbb{R}^n$, and satisfying the following standard growth, ellipticity/parabolicity and continuity conditions:

$$(1.18) \quad \begin{cases} |a(x, t, z)| + |a_z(x, t, z)|(|z| + s) \leq L(|z| + s) \\ \nu|\lambda|^2 \leq \langle a_z(x, t, z)\lambda, \lambda \rangle \\ |a(x, t, z) - a(x_0, t, z)| \leq L_1\omega(|x - x_0|)(|z| + s) \end{cases}$$

for every choice of $x, x_0 \in \Omega$, $z, \lambda \in \mathbb{R}^n$ and $t \in (-T, 0)$; here the function $\omega: [0, \infty) \rightarrow [0, 1]$ is as in (1.2)₃ for $p = 2$. Note that anyway we are assuming no continuity on the map $t \mapsto a(\cdot, t, \cdot)$, which is considered to be a priori only measurable. In other words we are considering the analog of assumptions (1.2) for $p = 2$; the reason we are adopting this restriction is that when dealing with the evolutionary p -Laplacean operator estimates take the usual form only when using so called “intrinsic cylinders”, according the by now classical parabolic p -Laplacean theory developed by DiBenedetto [17]. These are - unless $p = 2$ when they reduce to the standard parabolic ones - cylinders whose size locally depends on the size of the solutions itself, therefore a formulation of the estimates via non-linear potentials - whose definition is built essentially using a standard family of balls and it is therefore “universal” - is not clear and will be the object of future investigation. We refer to [2] for global gradient estimates.

In order to state our results we need some additional terminology. Let us recall that given points $(x, t), (x_0, t_0) \in \mathbb{R}^{n+1}$ the standard parabolic metric is defined by

$$(1.19) \quad d_{\text{par}}((x, t), (x_0, t_0)) := \max\{|x - x_0|, \sqrt{|t - t_0|}\} \approx \sqrt{|x - x_0|^2 + |t - t_0|}$$

and the related metric balls with radius R with respect to this metric are given by the cylinders of the type $B(x_0, R) \times (t_0 - R^2, t_0 + R^2)$. The “caloric” Riesz potential - compare with elliptic one defined in (1.11) and see Remark 1.2 below - is now built starting from

$$(1.20) \quad I_\beta(\mu)((x, t)) := \int_{\mathbb{R}^{n+1}} \frac{d\mu((\tilde{x}, \tilde{t}))}{d_{\text{par}}((\tilde{x}, \tilde{t}), (x, t))^{N-\beta}} \quad 0 < \beta \leq N := n + 2$$

whenever $(x, t) \in \mathbb{R}^{n+1}$. In order to be used in estimates for parabolic equations, it is convenient to introduce its local version via the usual backward parabolic cylinders - with “vertex” at (x_0, t_0) - that is

$$(1.21) \quad Q(x_0, t_0; R) := B(x_0, R) \times (t_0 - R^2, t_0),$$

and is now given by - with $N := n + 2$

$$(1.22) \quad \mathbf{I}_\beta^\mu(x_0, t_0; R) := \int_0^R \frac{|\mu|(Q(x_0, t_0; \varrho))}{\varrho^{N-\beta}} \frac{d\varrho}{\varrho} \quad \text{where } \beta \in (0, N].$$

The main result in the parabolic case is

Theorem 1.2 (Caloric potential gradient bound). *Under the assumptions (1.18) and (1.3), let $u \in C^0(-T, 0; L^2(\Omega))$ be a weak solution to (1.17) with $\mu \in L^2(\Omega_T)$, and such that $Du \in C^0(\Omega_T)$. Then there exists a constant $c \equiv c(n, \nu, L)$ and a radius $\tilde{R} \equiv \tilde{R}(n, \nu, L, L_1, \omega(\cdot))$ such that the following estimate:*

$$(1.23) \quad |Du(x_0, t_0)| \leq c \int_{Q(x_0, t_0; R)} (|Du| + s) dx dt + c \mathbf{I}_1^\mu(x_0, t_0; 2R)$$

holds whenever $Q(x_0, t_0; 2R) \subseteq \Omega$ is a backward parabolic cylinder with vertex at (x_0, t_0) and such that $R \leq \tilde{R}$.

As in the elliptic case when dealing with parabolic equations with no space variable dependence of the type

$$(1.24) \quad u_t - \operatorname{div} a(t, Du) = \mu$$

we can improve the result of Theorem 1.2 as follows:

Theorem 1.3 (Componentwise caloric bound). *Under the assumptions (1.18), let $u \in C^0(-T, 0; L^2(\Omega))$ be a weak solution to (1.24) with $\mu \in L^2(\Omega_T)$, and such that $Du \in C^0(\Omega_T)$. Then there exists a constant $c \equiv c(n, \nu, L)$ such that the following estimate:*

$$(1.25) \quad |D_\xi u(x_0, t_0)| \leq c \int_{Q(x_0, t_0; R)} (|D_\xi u| + s) dx dt + c \mathbf{I}_1^\mu(x_0, t_0; 2R)$$

holds whenever $Q(x_0, t_0; 2R) \subseteq \Omega$ is a backward parabolic cylinder with vertex at (x_0, t_0) , and $\xi \in \{1, \dots, n\}$.

Estimates (1.23)-(1.25) also hold for general weak and very weak solutions, and in particular for solutions to measure data problems as

$$(1.26) \quad \begin{cases} u_t - \operatorname{div} a(x, t, Du) = \mu & \text{in } \Omega_T \\ u = 0 & \text{on } \partial_{\text{par}} \Omega_T, \end{cases}$$

where μ is a general Radon measure on Ω_T with finite mass on Ω_T , that we shall consider to be defined in the whole \mathbb{R}^{n+1} . See next section for the definition of solutions. In the spirit of the elliptic result (1.9) we have the following implication, which provides a boundedness criteria for the spatial gradient, under the Dini-continuity assumption for the spatial coefficients (1.3):

$$(1.27) \quad \mathbf{I}_1^\mu(\cdot; R) \in L^\infty(\Omega_T), \text{ for some } R > 0 \implies Du \in L_{\text{loc}}^\infty(\Omega_T, \mathbb{R}^n).$$

We conclude with the zero order potential estimate, which applies to general equations of the type (1.17) when considered with a measurable dependence upon the coefficients (x, t) . The relevant hypotheses in this case are the following standard growth and monotonicity properties:

$$(1.28) \quad \begin{cases} |a(x, t, z)| \leq L(|z| + s) \\ \nu|z_2 - z_1|^2 \leq \langle a(x, t, z_2) - a(x, t, z_1), z_2 - z_1 \rangle \end{cases}$$

which are assumed to hold whenever $(x, t) \in \Omega_T$ and $z, z_1, z_2 \in \mathbb{R}^n$. In particular, since the pointwise bound will be derived on u , rather than on Du , we do not need differentiability assumptions for $a(\cdot)$ with respect to the spatial gradient variable z .

Theorem 1.4 (Zero order estimate). *Under the assumptions (1.28), let*

$$u \in L^2(-T, 0; W^{1,2}(\Omega)) \cap C^0(\Omega_T)$$

be a weak solution to (1.17) with $\mu \in L^2(\Omega_T)$. Then there exists a constant c , depending only on n, ν, L, L_1 , such that the following inequality holds whenever $Q(x_0, t_0; 2R) \subseteq \Omega$ is a backward parabolic cylinder with vertex at (x_0, t_0) :

$$(1.29) \quad |u(x_0, t_0)| \leq c \int_{Q(x_0, t_0; R)} (|u| + Rs) \, dx \, dt + c \mathbf{I}_2^\mu(x_0, t_0; 2R).$$

Remark 1.2. The caloric Riesz potential defined in (1.20) is different from the parabolic Riesz potential considered for instance in [3] via convolution of μ with the heat kernel

$$(1.30) \quad \Gamma_\beta(x, t) := \frac{1}{t^{\frac{N-\beta}{2}}} \exp\left(-\frac{|x|^2}{4t}\right) \quad \beta \in (0, N].$$

On the other hand both the kernels considered exhibit the same scaling with respect to the parabolic dilation $R \rightarrow (Rx, R^2t)$, which is in turn the relevant property to determine the regularization properties of the related convolution operator. Therefore, as we shall see from Section 6.1 below, we have that the caloric Riesz potential (1.20) is perfectly sufficient to infer the sharp regularity/integrability properties of solutions. In particular, from (1.23) we shall be able to derive a borderline Marcinkiewicz estimate which seems to be difficult to derive using for instance the truncation methods from [9]. Anyway, for estimates regarding (1.30) see [37].

1.4. Plan of the paper. After establishing in Section 2 some notation, Section 3 is dedicated to the proof of Theorem 1.1. This will require a careful combination of regularity estimates for p -harmonic functions, together with suitable comparison estimates which make the density of the Wolff potential appear; all such ingredients will be finally combined in a suitable iteration scheme. The same scheme will be followed in Section 4 for the parabolic case, where additional difficulties come into the play; in particular, the proof of the comparison estimates necessitates further delicate technicalities, while some precise estimates from Nash-Moser's theory will be needed. In Section 5 we show how the everywhere pointwise estimates derived for a priori regular solutions, actually extend to all kinds of general weak solutions, including solutions to measure data problems, which are the most general ones. Finally, in Section 6 we demonstrate how the pointwise estimates derived here allows to recast essentially all the main gradient estimates for non-homogeneous problems as described at the beginning of this Introduction, and in particular those in finer scales of spaces such as Lorentz or Orlicz spaces. We re-emphasize here that, when considering parabolic problems, the techniques presented here are the only one available for getting such estimates. The techniques presented in this paper are general enough to be applied in different contexts; an example is [12], where problems with non-standard growth of $p(x)$ -type - see for instance [1] - are considered.

Some of the results here have been announced in the *nota lincea* [22].

Acknowledgments. This research is supported by the ERC grant 207573 “Vectorial Problems”, and by MIUR via the national project “Calcolo delle Variazioni”. We also thank Verena Bögelein who carefully read a draft of the paper.

2. NOTATIONS

In this paper we follow the usual convention of denoting by c a general constant larger (or equal) than one, possibly varying from line to line; special occurrences

will be denoted by c_1 etc; relevant dependence on parameters will be emphasized using parentheses. We shall denote in a standard way

$$B(x_0, R) := \{x \in \mathbb{R}^n : |x - x_0| < R\}$$

the open ball with center x_0 and radius $R > 0$, while backward parabolic cylinders have been already defined in (1.21); when not important, or clear from the context, we shall omit denoting the center as follows: $B_R \equiv B(x_0, R)$, and the same will happen for parabolic cylinders concerning the vertex. Moreover, when more than one ball - resp. cylinder - will come into the play, they will always share the same center - resp. vertex - unless otherwise stated. We shall also denote $B \equiv B_1 = B(0, 1)$, and $Q \equiv Q_1 = B_1 \times (-1, 0)$. We recall that, given a cylindrical domain of the type $C = A \times (t_1, t_2)$ - and in particular a parabolic cylinder - its parabolic boundary $\partial_{\text{par}} C$ is given by $\partial C \setminus (A \times \{t_2\})$. With A being a measurable subset with positive measure, and with $g: A \rightarrow \mathbb{R}^k$ being a measurable map, we shall denote

$$\int_A g(x) dx := \frac{1}{|A|} \int_A g(x) dx$$

its average. In this paper, all the measures considered will be Radon measure with finite total mass, for convenience assumed to be defined on the whole space: \mathbb{R}^n for elliptic problems, \mathbb{R}^{n+1} for parabolic ones; with $A \subset \mathbb{R}^n$ or $A \subset \mathbb{R}^{n+1}$, we shall denote by $\mathcal{M}(A)$ the space of all Radon measures with finite total mass defined on A . When a measure μ actually turns out to be an L^1 -function, we shall use the standard notation, whenever A is a measurable set on which μ is defined

$$|\mu|(A) := \int_A |\mu(x)| dx.$$

When dealing with general measure data problems as in (1.1) and (1.17), we shall consider so called very weak solutions, i.e. those solutions who are not necessarily lying in the natural energy spaces associated to such equations growth assumptions (1.2)₁ and (1.18)₁ being given - specifically $W^{1,p}(\Omega)$ and $L^2(-T, 0; W^{1,2}(\Omega))$, respectively - but that are nevertheless integrable enough to allow for a usual distributional formulation. Such solutions are not necessarily unique, but suitable reinforced notions of solutions can be considered as well, in order to achieve uniqueness in several cases; see Section 5 below. In the elliptic case a (very) weak solution to (1.1) is a function $u \in W^{1,p-1}(\Omega)$ such that the distributional relation

$$\int_{\Omega} \langle a(x, Du), D\varphi \rangle dx = \int_{\Omega} \varphi d\mu$$

holds whenever $\varphi \in C_0^\infty(\Omega)$ has compact support in Ω . In the parabolic case a (very) weak solution is a function $u \in L^1(-T, 0; W^{1,1}(\Omega))$ such that

$$(2.1) \quad - \int_{\Omega_T} u \varphi_t dx dt + \int_{\Omega_T} \langle a(x, t, Du), D\varphi \rangle dx dt = \int_{\Omega_T} \varphi d\mu$$

holds whenever $\varphi \in C_0^\infty(\Omega_T)$ has compact support in Ω_T . When referring to Cauchy-Dirichlet problems of the type (1.26), while the lateral boundary condition can be formulated by prescribing the inclusion $u \in L^1(-T, 0; W_0^{1,1}(\Omega))$, the initial boundary one $u(x, -T) \equiv 0$ is understood in the L^1 -sense, that is prescribing that

$$\lim_{h \downarrow 0} \frac{1}{h} \int_{-T}^{-T+h} \int_{\Omega} |u(x, t)| dx dt = 0.$$

In the parabolic setting a convenient “slicewise” reformulation of (2.1) is given by mean of so-called Steklov averages; in fact, for $h > 0$ and $t \in [-T, 0)$ we define

$$(2.2) \quad u_h(x, t) := \begin{cases} \frac{1}{h} \int_t^{t+h} u(x, \tilde{t}) d\tilde{t} & \text{if } t+h < 0 \\ 0 & \text{if } t+h > 0. \end{cases}$$

With such a notation we have - see [17, Chapter 2] - that when $\mu \in L^1(\Omega_T)$ the slicewise equality

$$(2.3) \quad \int_{\Omega} \left(\partial_t u_h \varphi + \langle [a(\cdot, t), Du]_h, D\varphi \rangle \right) dx = \int_B \varphi \mu_h dx$$

holds whenever $\varphi \in C_0^\infty(\Omega)$ has compact support in Ω , and for a.e. $t \in (-T, 0)$. As the reader will recognize we use both the notations w_t and $\partial_t w$ for the time derivative of a function w .

3. ELLIPTIC ESTIMATES

This section contains the proof of estimate (1.6), therefore we shall argue under the assumptions of Theorem 1.1, while $u \in C^1(\Omega)$ denotes the solution identified in Theorem 1.1.

3.1. Basic preliminaries. Let us recall a few basic strict monotonicity properties of the vector field $z \mapsto a(\cdot, z)$ under the assumption (1.2)₂; for an absolute constant $c \equiv c(n, p, \nu, L) \geq 1$ the inequality With $s \geq 0$ being the number appearing in (1.2), we define

$$(3.1) \quad c^{-1}(|z_1|^2 + |z_2|^2 + s^2)^{\frac{p-2}{2}} |z_2 - z_1|^2 \leq \langle a(x, z_2) - a(x, z_1), z_2 - z_1 \rangle$$

holds whenever z_1, z_2, \mathbb{R}^n and $x \in \Omega$. In particular since $p \geq 2$ the previous inequality implies Indeed using (1.2)₂ is standard to see that the inequality

$$(3.2) \quad c^{-1} |z_2 - z_1|^p \leq \langle a(x, z_2) - a(x, z_1), z_2 - z_1 \rangle.$$

The next result on Reverse Hölder type inequalities will be useful in the sequel.

Lemma 3.1. *Let $g : A \rightarrow \mathbb{R}^k$ be a integrable map such that*

$$\left(\int_{B_R} |g|^{\chi_0} dx \right)^{\frac{1}{\chi_0}} \leq c \int_{B_{2R}} |g| dx$$

holds whenever $B_{2R} \subseteq A$, where $A \subseteq \mathbb{R}^n$ is an open subset, and $\chi_0 > 1$, $c \geq 0$. Then, for every $t \in (0, 1]$ and $\chi \in (0, \chi_0]$ there exists a constant $c_0 \equiv c_0(n, c, t)$ such that, for every $B_{2R} \subseteq A$ it holds that

$$(3.3) \quad \left(\int_{B_R} |g|^\chi dx \right)^{\frac{1}{\chi}} \leq c_0 \left(\int_{B_{2R}} |g|^t dx \right)^{\frac{1}{t}}.$$

The assertion also holds when $A \subseteq \mathbb{R}^{n+1}$, considering backward parabolic cylinders of the type in (1.21), instead of standard balls.

3.2. A decay estimate below the natural growth exponent. Here we prove a decay estimate for solutions to certain homogeneous equations which differs from more standard ones in that the exponents involved are smaller than those typically used - the “natural growth exponents”; this estimate could be of its own interest. Let us consider an energy solution $v \in W^{1,p}(A)$ to the homogeneous equation

$$(3.4) \quad \operatorname{div} a(x_0, Dv) = 0,$$

where $x_0 \in \Omega$ and $A \subseteq \Omega$ is a sub-domain of Ω , in other words we consider the vector field from Theorem 1.1 “frozen” at a point.

Lemma 3.2. *Let $w \in W^{1,2}(A)$ be a solution of the linear equation of the type*

$$\operatorname{div}(\tilde{a}(x)Dw) = 0,$$

where the matrix $\tilde{a}(x)$ has measurable entries and satisfies the following ellipticity and growth bounds, for every $\lambda \in \mathbb{R}^n$:

$$c_*|\lambda|^2 \leq \langle \tilde{a}(x)\lambda, \lambda \rangle, \quad |\tilde{a}(x)| \leq c_{**}, \quad \text{for some } c_* \in (0, 1) \text{ and } c_{**} \geq 1.$$

*Then there exists constants $c \geq 1$ and $\beta_0 \in (0, 1]$, both depending only on $n, c_{**}/c_*$, such that the following estimate holds whenever $B_\varrho \subset B_R \subseteq A$ are concentric balls:*

$$(3.5) \quad \int_{B_\varrho} |w - (w)_{B_\varrho}| dx \leq c \left(\frac{\varrho}{R} \right)^{\beta_0} \int_{B_R} |w - (w)_{B_R}| dx.$$

Proof. This result is a rather standard consequence of DeGiorgi's theory for linear elliptic equations. The statement follows directly from [39, Corollary 1.5] or by the proof of [24, Theorem 7.7], see also [24, (7.45)], taking into account the specific equations we are dealing with; in particular we may take $\chi = 0$ in [24]. \square

Theorem 3.1. *Let $v \in W^{1,p}(A)$ be a weak solution to (3.4) under the assumptions (1.2). Then there exist constants $\beta \in (0, 1]$ and $c \geq 1$, both depending only on n, p, ν, L , such that the estimate*

$$(3.6) \quad \int_{B_\varrho} |Dv - (Dv)_{B_\varrho}| dx \leq c \left(\frac{\varrho}{R} \right)^\beta \int_{B_R} |Dv - (Dv)_{B_R}| dx$$

holds whenever $B_\varrho \subseteq B_R \subseteq A$ are concentric balls.

Proof. We shall use some iteration techniques [18, 39, 21]; in particular, we shall use the ideas in the important paper [39]. The proof is divided in several steps. In the following $B(x_0, R) \equiv B_R$ is a fixed ball as considered in the statement of the Theorem. We recall that standard regularity gives that $Dv \in L^\infty_{\text{loc}}(A)$; moreover without loss of generality we can assume that $B_R \Subset A$ so that in the following we have that $Dv \in L^\infty(B_R)$. From now on, for the rest of the proof, all the balls considered will have the same center x_0 and r will denote a positive radius such that $r \leq R$. Accordingly, we set

$$E(r) := \int_{B_r} |Dv - (Dv)_{B_r}| dx, \quad M(r) := \max_{1 \leq \xi \leq n} \sup_{B_r} |D_\xi v|.$$

Step 1: Regularization. In the following we shall consider the non-degenerate case $s > 0$, but we shall prove estimates which do not depend on the parameter s . The case $s = 0$ can be deduced by approximating original solutions by solutions to equations of similar type satisfying $s > 0$, via a completely standard approximation procedure; see for instance [16, 46].

Step 2: The fundamental alternative. This goes as follows - see [16, 46, 39]: With $B_{2r} \subseteq B_R$ and $\mu_0 \in (0, 1)$, clearly one of the following three alternatives must hold:

$$(3.7) \quad \left\{ \begin{array}{ll} |\{D_\xi v < M(2r)/2\} \cap B_{2r}| \leq \mu_0 |B_{2r}| & \text{for some } \xi \\ |\{D_\xi v > -M(2r)/2\} \cap B_{2r}| \leq \mu_0 |B_{2r}| & \text{for some } \xi \\ |\{D_\xi v < M(2r)/2\} \cap B_{2r}|, \\ |\{D_\xi v > -M(2r)/2\} \cap B_{2r}| > \mu_0 |B_{2r}| & \text{for all } \xi. \end{array} \right.$$

The crucial and well-established regularity property of solutions - see for instance [16, 39, 46] - is that there exist *universal numbers* $\mu_0, \eta \equiv \mu_0, \eta(n, p, \nu, L) \in (0, 1)$ - i.e. independent of the vector field considered $a(x_0, z)$ and of the solution v - such that if one of the first two alternatives in (3.7) holds then

$$(3.8) \quad |Dv| \geq M(2r)/4 \quad \text{in } B_r$$

otherwise, if the third possibility from (3.7) occurs, we have that

$$(3.9) \quad M(r) \leq \eta M(2r).$$

All in all we have that either (3.8) or (3.9) happens. The proof now consists of showing that combining these two alternatives (3.6) follows. We recall that under the present assumptions we have that $v \in W_{\text{loc}}^{2,2}(A)$ and each component $D_\xi v \equiv w$ satisfies the following differentiated equation:

$$(3.10) \quad \operatorname{div}(\tilde{a}(x)Dw) = 0 \quad \tilde{a}_{i,j}(x) := (a_i)_{z_j}(x_0, Dv(x)).$$

Moreover, standard a priori estimates for p -Laplacean type equations as those considered in (3.4) provide a constant $c_g \equiv c_g(n, p, \nu, L)$ such that

$$(3.11) \quad \sup_{B_r} |Dv| \leq c_g \int_{B_{2r}} (|Dv| + s) dx \quad \text{holds for every } B_{2r} \subseteq B_R.$$

Step 3: The non-degenerate case I. In this case we assume that $s > 2\sqrt{n}M(r)$ for a certain radius $r \leq R$; then the matrix $\tilde{a}(x)$ in (3.10) satisfies on B_r the bounds

$$\nu^{-1}s^{p-2}|\lambda|^2 \leq \langle \tilde{a}(x)\lambda, \lambda \rangle, \quad |\tilde{a}(x)| \leq L2^{p-2}s^{p-2} \quad \text{for every } \lambda \in \mathbb{R}^n.$$

Therefore the matrix $\tilde{a}(x)$ defined in (3.10) satisfies the assumptions of Lemma 3.2, that applied to every component $D_\xi v \equiv w$ of the gradient gives

$$(3.12) \quad \int_{B_\rho} |Dv - (Dv)_{B_\rho}| dx \leq c_d \left(\frac{\rho}{r}\right)^{\beta_0} \int_{B_r} |Dv - (Dv)_{B_r}| dx$$

where $c_d \geq 1, \beta_0 \in (0, 1)$ depend only n, p, ν, L .

Step 4: Choice of the constants. We first take a positive $H_1 \in \mathbb{N}$ such that

$$(3.13) \quad 8\sqrt{n}c_g\eta^{H_1-1} \leq 1.$$

This fixes H_1 as a quantity depending only on c_g, η and therefore ultimately on n, p, ν, L . In turn we determine another integer K_1 such that

$$(3.14) \quad 2^{nH_1+2}\eta^{K_1} \leq 1, \quad c_d 2^{-K_1\beta_0+nH_1+2} \leq 1$$

hold. This fixes K_1 as a quantity depending on η, H_1, c_d , and therefore ultimately on n, p, ν, L .

Step 5: The degenerate case I. We consider the following situation: there exists a radius $t \leq R$ such that (3.9) happens to hold whenever $r = t/2^i$ and $1 \leq i \leq H_1 \in \mathbb{N}$, and finally we assume that both

$$(3.15) \quad |(Dv)_{B_t}| \leq 2\sqrt{n}M(2^{-H_1}t) \quad \text{and} \quad s \leq 2\sqrt{n}M(2^{-H_1}t)$$

hold. We first notice that $E(2^{-H_1}t) \leq 2\sqrt{n}M(2^{-H_1}t)$, while iterating (3.9) yields

$$M(2^{-H_1}t) \leq \eta^{H_1-1}M(t/2).$$

On the other hand we have that

$$\begin{aligned} M(2^{-H_1}t) &\leq \eta^{H_1-1}M(t/2) \\ &\leq c_g\eta^{H_1-1} \int_{B_t} (|Dv| + s) dx \\ &\leq c_g\eta^{H_1-1} \int_{B_t} |Dv - (Dv)_{B_t}| dx + c_g\eta^{H_1-1} [| (Dv)_{B_t} | + s] \\ &\leq c_g\eta^{H_1-1} E(t) + 4\sqrt{n}c_g\eta^{H_1-1} M(2^{-H_1}t), \end{aligned}$$

where we used (3.11). Therefore we have $M(2^{-H_1}t) \leq 2\eta^{H_1-1}c_gE(t)$, so that, again by (3.13), the final outcome is

$$(3.16) \quad E(2^{-H_1}t) \leq \frac{1}{2}E(t).$$

Step 6: The degenerate case II. Continuing the reasoning of the previous step, we assume that there exists a radius $t \leq R$ such that (3.9) happens to hold whenever $r = t/2^i$ and $1 \leq i \leq H_1 + K_1 \in \mathbb{N}$, assuming also, in alternative to (3.15), that *at least* one of the following inequalities holds:

$$(3.17) \quad |(Dv)_{B_t}| > 2\sqrt{n}M(2^{-H_1}t) \quad \text{and} \quad s > 2\sqrt{n}M(2^{-H_1}t).$$

In the case the first inequality in (3.17) holds we have that also

$$|Dv - (Dv)_{B_t}| > \sqrt{n}M(2^{-H_1}t) \quad \text{in } B_{2^{-H_1}t};$$

therefore, using also (3.14), we conclude as in (3.16):

$$\begin{aligned} E(2^{-(H_1+K_1)}t) &\leq 2\sqrt{n}M(2^{-(H_1+K_1)}t) \\ &\leq 2\sqrt{n}\eta^{K_1}M(2^{-H_1}t) \\ &\leq 2\eta^{K_1} \int_{B_{2^{-H_1}t}} |Dv - (Dv)_{B_t}| dx \\ &\leq 2^{nH_1+1} \eta^{K_1} \int_{B_t} |Dv - (Dv)_{B_t}| dx \\ (3.18) \quad &\leq \frac{1}{2}E(t). \end{aligned}$$

On the other hand, if $s > 2\sqrt{n}M(2^{-H_1}t)$, we are in the situation of Step 3 with the choice $r \equiv 2^{-H_1}t$ and $\varrho \equiv 2^{-(H_1+K_1)}t$; therefore applying (3.12) yields

$$E(2^{-(H_1+K_1)}t) \leq c_d 2^{-K_1\beta_0} E(2^{-H_1}t)$$

while it also holds

$$E(2^{-H_1}t) \leq 2 \int_{B_{2^{-H_1}t}} |Dv - (Dv)_{B_t}| dx \leq 2^{nH_1+1} E(t),$$

so that, merging the last two estimates and using (3.14) we again conclude as in (3.18) with

$$E(2^{-(H_1+K_1)}t) \leq c_d 2^{-K_1\beta_0+nH_1+1} E(t) \leq \frac{1}{2}E(t).$$

Step 7: Summarizing Step 5 and Step 6. Here we first note that the result of Step 3 holds if we replace H_1 by any larger integer since if (3.13) holds for H_1 , then so it does for any larger number. Therefore, looking at Step 6; if (3.9) holds for $r = t/2^i$ and $1 \leq i \leq H_1 + K_1 \in \mathbb{N}$, and assuming also both

$$|(Dv)_{B_t}| \leq 2\sqrt{n}M(2^{-H_1-K_1}t) \quad \text{and} \quad s < 2\sqrt{n}M(2^{-H_1-K_1}t),$$

we can apply Step 5 with H_1 replaced by $H_1 + K_1$. Therefore, summarizing the results of Steps 5 and 6 we have that: If (3.9) holds for $r = t/2^i$ and $1 \leq i \leq H := H_1 + K_1$ then

$$(3.19) \quad E(\tau t) \leq \frac{1}{2}E(t), \quad \tau := 1/2^H.$$

The crucial fact is that H , and therefore also τ , only depends on n, p, ν, L .

Step 8: Conclusion. Let us define $\beta_1 := 1/H$ so that $\beta_1 \equiv \beta_1(n, p, \nu, L) \in (0, 1]$ and $\tau^{\beta_1} = 1/2$, where $\tau \in (0, 1/2]$ has been introduced in (3.19); we finally determine the exponent β appearing in (3.6) by letting $\beta := \min\{\beta_0, \beta_1\}$ where β_0 is the exponent from Lemma 3.2. Moreover, with $\varrho \leq R$ as in the statement of the Theorem, we fix $k \in \mathbb{N}$ such that $\tau^{k+1}R < \varrho \leq \tau^k R$. Let us now define

$$\mathbb{S} := \{i \in \mathbb{N} : (3.9) \text{ holds for } r = R/2^i, i \geq 1\}.$$

We argue on an alternative; the first case is when $\mathbb{S} = \mathbb{N} \setminus \{0\}$. Then either Step 5 or Step 6, and so Step 7, apply for every choice of the radius $t \equiv \tau^i R$ and $i \geq 0$, thus obtaining that

$$(3.20) \quad E(\tau^i R) \leq (1/2)^i E(R) = \tau^{i\beta_1} E(R) \leq \tau^{i\beta} E(R)$$

holds for every $i \in \mathbb{N}$. Taking into account the definition of β_1 we have

$$E(\varrho) \leq 2\tau^{-n} E(\tau^k R) \leq c\tau^{k\beta} E(R) \leq c\left(\frac{\varrho}{R}\right)^\beta E(R),$$

so that (3.6) follows; note here that $c \equiv c(\tau) \equiv c(n, p, \nu, L)$ by (3.19). In the other case we have that $\mathbb{S} \neq (\mathbb{N} \setminus \{0\})$, and therefore if we set $m := \min((\mathbb{N} \setminus \{0\}) \setminus \mathbb{S})$ then (3.8) implies that $|Dv| \geq M(R/2^{m-1})$ in the ball $B_{R/2^m}$. At this point we observe that in the ball $B_{R/2^m}$ the matrix $\tilde{a}(x)$ defined in (3.10) satisfies the following uniform ellipticity bounds:

$$c^{-1}[M(R/2^{m-1}) + s]^{p-2}|\lambda|^2 \leq \langle \tilde{a}(x)\lambda, \lambda \rangle, \quad |\tilde{a}(x)| \leq c[M(R/2^{m-1}) + s]^{p-2}$$

for $c \equiv c(n, p, \nu, L) > 1$ and every choice of $\lambda \in \mathbb{R}^n$, and therefore we may apply Lemma 3.2 as already done in Step 3, thereby getting

$$(3.21) \quad E(r) \leq c\left(\frac{r}{R/2^m}\right)^{\beta_0} E(R/2^m) \quad \text{for every } r \leq R/2^m.$$

Let now $\gamma \in \mathbb{N}$ be the unique non-negative integer such that $\gamma H < m \leq (\gamma + 1)H$. We apply Step 7 exactly γ times thereby getting - also when $\gamma = 0$ - that

$$(3.22) \quad E(\tau^i R) \leq (1/2)^i E(R) \quad \text{for every } i \leq \gamma.$$

Moreover, using the definition of $E(\cdot)$ and τ it immediately follows that

$$E(R/2^m) \leq 2 \int_{B_{R/2^m}} |Dv - (Dv)_{B_{\tau^\gamma R}}| dx \leq 2^{nH+1} E(\tau^\gamma R),$$

so that combining the last two inequalities gives

$$E(R/2^m) \leq 2^{nH+1} 2^{-\gamma} E(R) = c\tau^{\gamma\beta_1} E(R) \leq c\tau^{\gamma\beta} E(R)$$

with $c \equiv c(n, p, \nu, L)$. On the other hand using (3.21) with $r = \tau^{\gamma+l} R \equiv 2^{-(\gamma+l)H} R$ and $l \in (\mathbb{N} \setminus \{0\})$, we gain

$$\begin{aligned} E(\tau^{\gamma+l} R) &\leq c\left(\frac{\tau^{\gamma+l} R}{R/2^m}\right)^{\beta_0} E(R/2^m) \\ &\leq c 2^{H\beta_0} \tau^{\ell\beta_0} E(R/2^m) \\ &\leq c \tau^{(\gamma+l)\beta} E(R). \end{aligned}$$

The last estimate and (3.22) in turn imply, for a suitable $c \equiv c(n, p, \nu, L)$, that the inequality $E(\tau^i R) \leq c\tau^{i\beta} E(R)$ holds for every $i \in \mathbb{N}$. At this point we again conclude as after (3.20). \square

3.3. Comparison estimates. This section is devoted to the proof of a few suitable comparison estimates in which the density of the non-linear Wolff potential $\mathbf{W}_{\frac{1}{p}, p}^\mu$ explicitly comes into the play. We now fix, for the rest of Section 3.3, a ball $B_{2R} \equiv B(x_0, 2R) \subseteq \Omega$ with radius $2R$; we start defining $w \in u + W_0^{1,p}(B_{2R})$ as the unique solution to the homogeneous Dirichlet problem

$$(3.23) \quad \begin{cases} \operatorname{div} a(x, Dw) = 0 & \text{in } B_{2R} \\ w = u & \text{on } \partial B_{2R}. \end{cases}$$

Lemma 3.3. *Under the assumption (1.2)₂, let $u \in W^{1,p}(\Omega)$ be as in Theorem 1.1, and $w \in u + W_0^{1,p}(B_{2R})$ as in (3.23). Then the following inequalities hold for a constant $c \equiv c(n, p, \nu)$:*

$$(3.24) \quad \int_{B_{2R}} |Du - Dw| dx \leq c \left[\frac{|\mu|(B_{2R})}{R^{n-1}} \right]^{\frac{1}{p-1}}$$

$$(3.25) \quad \int_{B_{2R}} |u - w| dx \leq c \left[\frac{|\mu|(B_{2R})}{R^{n-p}} \right]^{\frac{1}{p-1}}.$$

Proof. The proof revisits and modifies various comparison and truncation methods for measure data problems scattered in the literature; a chief reference here is the work of Boccardo & Gall  uet [10]. The proof is divided in three steps. We may without loss of generality assume that $2R = 1$, that the ball in question is centered at the origin $B_{2R} \equiv B_1$, and finally that $|\mu|(B_1) = 1$; the scaling technique necessary to reduce to such a situation is reported in Step 3 below.

Step 1: The case $p > n$. We test the weak formulations of (1.1) and (3.23)₁

$$(3.26) \quad \int_{B_1} \langle a(x, Du) - a(x, Dw), D\varphi \rangle dx = \int_{B_1} \varphi d\mu,$$

with $\varphi \equiv u - w$. Using Morrey-Sobolev's embedding theorem we estimate the resulting right hand side as follows:

$$\begin{aligned} \left| \int_{B_1} (u - w) d\mu \right| &\leq \sup_{B_1} |u - w| [|\mu|(B_1)] \\ &\leq \sup_{B_1} |u - w| \leq \frac{cp}{p-n} \left(\int_{B_1} |Du - Dw|^p dx \right)^{\frac{1}{p}}. \end{aligned}$$

The last inequality used together with (3.26) with $\varphi = u - w$ and (3.2) yields

$$\int_{B_1} |Du - Dw|^p dx \leq c \left(\int_{B_1} |Du - Dw|^p dx \right)^{\frac{1}{p}},$$

so that we first get $\|Du - Dw\|_{L^p(B_1)} \leq c$ and eventually $\|Du - Dw\|_{L^1(B_1)} \leq c$ that is (3.24) - when $B_{2R} \equiv B_1$ and $|\mu|(B_1) = 1$.

Step 2: The case $2 \leq p \leq n$. For $k \geq 0$ denoting an integer, we define the following truncation operators

$$(3.27) \quad T_k(s) := \max\{-k, \min\{s, k\}\} \quad \text{and} \quad \Phi_k(s) := T_1(s - T_k(s)),$$

defined for $s \in \mathbb{R}$. We test the weak formulation (3.26) by $\varphi \equiv T_k(u - w)$. By (3.2), setting $D_k := \{x \in B_1 : |u(x) - w(x)| \leq k\}$, we obtain, with $c \equiv c(n, p, \nu)$,

$$(3.28) \quad \int_{D_k} |Du - Dw|^p dx \leq ck.$$

Similarly, testing (3.26) with $\varphi \equiv \Phi_k(u - w)$ yields

$$(3.29) \quad \int_{C_k} |Du - Dw|^p dx \leq c,$$

where this time $C_k := \{x \in B_1 : k < |u(x) - w(x)| \leq k + 1\}$ and $c \equiv c(n, p, \nu)$. By H  lder's inequality, (3.29) and the very definition of C_k , we find

$$\begin{aligned} \int_{C_k} |Du - Dw| dx &\leq |C_k|^{\frac{p-1}{p}} \left(\int_{C_k} |Du - Dw|^p dx \right)^{\frac{1}{p}} \\ (3.30) \quad &\leq c |C_k|^{\frac{p-1}{p}} \leq ck^{-\frac{n(p-1)}{p(n-1)}} \left(\int_{C_k} |u - w|^{\frac{n}{n-1}} dx \right)^{\frac{p-1}{p}}. \end{aligned}$$

We have of course used the elementary estimate

$$|C_k| \leq \frac{1}{k^{\frac{n}{n-1}}} \int_{C_k} |u - w|^{\frac{n}{n-1}} dx.$$

Using (3.30) with (3.28), Hölder's inequality for sequences, and finally Sobolev's embedding theorem, with k_0 being a fixed positive integer we have

$$\begin{aligned}
\int_{B_1} |Du - Dw| dx &= \int_{D_{k_0}} (\dots) dx + \sum_{k=k_0}^{\infty} \int_{C_k} (\dots) dx \\
&\leq ck_0^{\frac{1}{p}} + c \sum_{k=k_0}^{\infty} k^{-\frac{n(p-1)}{p(n-1)}} \left(\int_{C_k} |u - w|^{\frac{n}{n-1}} dx \right)^{\frac{p-1}{p}} \\
&\leq ck_0^{\frac{1}{p}} + c \left[\sum_{k=k_0}^{\infty} \frac{1}{k^{\frac{n(p-1)}{n-1}}} \right]^{\frac{1}{p}} \left(\int_{B_1} |u - w|^{\frac{n}{n-1}} dx \right)^{\frac{p-1}{p}} \\
(3.31) \quad &\leq ck_0 + cH(k_0) \left(\int_{B_1} |Du - Dw| dx \right)^{\frac{(p-1)n}{(n-1)p}}.
\end{aligned}$$

In the last lines we have obviously set

$$H(k_0) := \left[\sum_{k=k_0}^{\infty} \frac{1}{k^{\frac{n(p-1)}{n-1}}} \right]^{\frac{1}{p}},$$

while c depends on n, p, ν ; note that

$$p \geq 2 \implies \frac{n(p-1)}{n-1} > 1$$

so that $H(k_0)$ is always finite and satisfies $H(k_0) \rightarrow 0$ when $k_0 \rightarrow \infty$. Now, if $p < n$ then

$$\frac{(p-1)n}{(n-1)p} < 1;$$

therefore we take $k_0 = 1$ in (3.31) and applying Young's inequality in (3.31) we find

$$(3.32) \quad \|Du - Dw\|_{L^1(B_1)} \leq c,$$

that is (3.24) when $2R = 1$ and $|\mu|(B_1) = 1$. When $p = n$ we choose $k_0 \equiv k_0(n, p, \nu)$ large enough in order to have $cH(k_0) = 1/2$ in (3.31) and reabsorb the last integral on the right hand side, so that (3.32) follows again.

Step 3: Scaling procedures. We first reduce to the case $B_{2R} \equiv B_1$ by a standard scaling argument, i.e. letting

$$\tilde{u}(y) := \frac{u(x_0 + 2Ry)}{2R}, \quad \tilde{w}(y) := \frac{w(x_0 + 2Ry)}{2R},$$

and

$$\tilde{a}(y, z) := a(x_0 + 2Ry, z), \quad \tilde{\mu}(y) := 2R\mu(x_0 + 2Ry)$$

for $y \in B_1$ so that $-\operatorname{div} \tilde{a}(y, D\tilde{u}) = \tilde{\mu}$ and $\operatorname{div} \tilde{a}(y, D\tilde{w}) = 0$ hold. At this point one writes estimate (3.24) for \tilde{u} and \tilde{w} and then scales back. To reduce to the case $|\mu|(B_1) = 1$ we adopt another scaling: this time we define $A := [|\mu|(B_1)]^{1/(p-1)}$, and we may assume $A > 0$ otherwise $u \equiv w$ and (3.24) follows trivially by the strict monotonicity of the operator $a(\cdot)$. We define the new solutions $\bar{u} := A^{-1}u$, $\bar{w} := A^{-1}w$, the new datum $\bar{\mu} := A^{1-p}\mu$, and the new vector field $\bar{a}(x, z) := A^{1-p}a(x, Az)$. Therefore we have $\operatorname{div} \bar{a}(x, D\bar{u}) = \bar{\mu}$, $\operatorname{div} \bar{a}(x, D\bar{w}) = 0$, in the weak sense. We make sure that we can apply the result in Step 2. Trivially $|\bar{\mu}|(B_1) = 1$ and moreover it is easy to see that the vector field $\bar{a}(x, z)$ satisfies (1.2) with s replaced by $s/A \geq 0$. Therefore (3.24) holds in the form

$$\int_{B_1} |D\bar{u} - D\bar{w}| dx \leq c_2$$

with $c_2 \equiv c_2(n, p, \nu)$. At this point we find back (3.24) just using the definitions of \bar{u}, \bar{w} . As for (3.25), this is a consequence of (3.24) via the use of Sobolev inequality:

$$\int_{B_{2R}} |u - w| dx \leq cR \int_{B_{2R}} |Du - Dw| dx \leq c \left[\frac{|\mu|(B_{2R})}{R^{n-p}} \right]^{\frac{1}{p-1}}.$$

The proof is now complete \square

Remark 3.1. In the proof of the gradient estimate (1.6) we shall only need (3.24), while (3.25) will be needed only later, in the different context of Remark 4.1 below.

After introducing w in (3.23) we similarly define $v \in w + W_0^{1,p}(B_R)$, on the concentric smaller ball $B_R \equiv B(x_0, R)$, as the unique solution to the homogeneous Dirichlet problem with frozen coefficients

$$(3.33) \quad \begin{cases} \operatorname{div} a(x_0, Dv) = 0 & \text{in } B_R \\ v = w & \text{on } \partial B_R. \end{cases}$$

Next, we state a standard comparison result; we report its proof for completeness.

Lemma 3.4. *Under the assumptions of Theorem 1.1, with w as in (3.23) and v as in (3.33), there exists a constant $c \equiv c(n, p, \nu, L)$ such that the following inequality holds:*

$$(3.34) \quad \int_{B_R} |Dw - Dv|^p dx \leq c[L_1\omega(R)]^2 \int_{B_R} (|Dw| + s)^p dx.$$

Proof. We note that by the growth condition on a in (1.2)₁ and by (3.2), we also have the following coercivity condition:

$$(3.35) \quad c^{-1}(|z|^2 + s^2)^{\frac{p-2}{2}}|z|^2 - cs^p \leq \langle a(x, z), z \rangle, \quad c \equiv c(n, p, \nu, L) \geq 1,$$

which easily follows by Young's inequality. Using (3.35), that is testing (3.33)₁ with $w - v$, and then (1.2)₂ and Young's inequality, gives the energy bound

$$(3.36) \quad \int_{B_R} |Dv|^p dx \leq c \int_{B_R} (|Dw| + s)^p dx,$$

which holds for a constant c depending on n, p, ν, L . In turn, using (3.1), the fact that both v and w are solutions, (1.2)₃ and again Young's inequality, we have

$$\begin{aligned} & \int_{B_R} (s^2 + |Dv|^2 + |Dw|^2)^{\frac{p-2}{2}} |Dw - Dv|^2 dx \\ & \leq c \int_{B_R} \langle a(x_0, Dw) - a(x_0, Dv), Dw - Dv \rangle dx \\ & = c \int_{B_R} \langle a(x_0, Dw) - a(x, Dw), Dw - Dv \rangle dx \\ & \leq cL_1\omega(R) \int_{B_R} (|Dv|^2 + |Dw|^2 + s^2)^{\frac{p-1}{2}} |Dw - Dv| dx \\ & \leq \frac{1}{2} \int_{B_R} (|Dv|^2 + |Dw|^2 + s^2)^{\frac{p-2}{2}} |Dw - Dv|^2 dx \\ & \quad + c[L_1\omega(R)]^2 \int_{B_R} (|Dv| + |Dw| + s)^p dx. \end{aligned}$$

Therefore we gain

$$\int_{B_{\bar{R}}} (s^2 + |Dv|^2 + |Dw|^2)^{\frac{p-2}{2}} |Dw - Dv|^2 dx \leq c[L_1\omega(R)]^2 \int_{B_R} (|Dv| + |Dw| + s)^p dx,$$

and (3.34) follows by last inequality recalling that $p \geq 2$. \square

Lemma 3.5. *Let u be as in Theorem 1.1, and let $w \in u + W_0^{1,p}(B_{2R})$ and $v \in w + W_0^{1,p}(B_R)$ be defined in (3.23) and (3.33), respectively. Then for a constant c depending only on n, p, ν, L , it holds that*

$$(3.37) \quad \begin{aligned} \int_{B_R} |Du - Dv| dx &\leq c \left\{ 1 + [L_1 \omega(R)]^{\frac{2}{p}} \right\} \left[\frac{|\mu|(B_{2R})}{R^{n-1}} \right]^{\frac{1}{p-1}} \\ &\quad + c [L_1 \omega(R)]^{\frac{2}{p}} \int_{B_{2R}} (|Du| + s) dx. \end{aligned}$$

Proof. We start proving the following inequality:

$$(3.38) \quad \int_{B_R} |Dw - Dv| dx \leq c [L_1 \omega(R)]^{\frac{2}{p}} \int_{B_{2R}} (|Dw| + s) dx,$$

with $c \equiv c(n, p, \nu, L)$. Keeping (3.35) in mind we may apply Gehring's lemma in the version presented in [24, Chapter 6], finding there exists a constant $\chi_0 \equiv \chi_0(n, p, \nu, L) > 1$ such that the reverse Hölder type inequality

$$\left(\int_{B_{\varrho/2}} (|Dw| + s)^{\chi_0 p} dx \right)^{\frac{1}{\chi_0}} \leq c \int_{B_{\varrho}} (|Dw| + s)^p dx$$

holds whenever $B_{\varrho} \subseteq B_{2R}$, for a constant c depending only on n, p, ν, L . In turn, applying Lemma 3.1 with $g \equiv (|Dw| + s)^p$, leads to establish that also

$$(3.39) \quad \left(\int_{B_R} (|Dw| + s)^p dx \right)^{\frac{1}{p}} \leq c \int_{B_{2R}} (|Dw| + s) dx$$

holds. Using now (3.34) and previous inequality we estimate as follows:

$$\begin{aligned} \int_{B_R} |Dw - Dv| dx &\leq c \left(\int_{B_R} |Dw - Dv|^p dx \right)^{\frac{1}{p}} \\ &\leq c [L_1 \omega(R)]^{\frac{2}{p}} \left(\int_{B_R} (|Dw| + s)^p dx \right)^{\frac{1}{p}} \leq c [L_1 \omega(R)]^{\frac{2}{p}} \int_{B_{2R}} (|Dw| + s) dx \end{aligned}$$

and therefore (3.38) follows. Using (3.38) and (3.24) now yields

$$(3.40) \quad \int_{B_R} |Du - Dv| dx \leq c \left[\frac{|\mu|(B_{2R})}{R^{n-1}} \right]^{\frac{1}{p-1}} + c [L_1 \omega(R)]^{\frac{2}{p}} \int_{B_{2R}} (|Dw| + s) dx.$$

In order to estimate the last integral in (3.40) we simply use (3.24) as follows:

$$\begin{aligned} \int_{B_{2R}} |Dw| dx &\leq \int_{B_{2R}} |Du| dx + \int_{B_{2R}} |Du - Dw| dx \\ &\leq \int_{B_{2R}} |Du| dx + c(n, p, \nu) \left[\frac{|\mu|(B_{2R})}{R^{n-1}} \right]^{\frac{1}{p-1}}. \end{aligned}$$

Using the last inequality in combination with (3.40) we conclude with (3.37). \square

3.4. Proof of the main estimates. We start with a technical lemma.

Lemma 3.6. *Let u be as in Theorem 1.1, then there exists constants $\beta \in (0, 1]$ and $c, c_1 \geq 1$, all depending only on n, p, ν, L , and a positive radius $R_1 \equiv R_1(L_1, \omega(\cdot))$, such that the following estimate holds whenever $B_{\varrho} \subseteq B_R \subseteq B_{2R} \subseteq \Omega$ are concentric balls with $R \leq R_1$:*

$$(3.41) \quad \begin{aligned} \int_{B_{\varrho}} |Du - (Du)_{B_{\varrho}}| dx &\leq c_1 \left(\frac{\varrho}{R} \right)^{\beta} \int_{B_{2R}} |Du - (Du)_{B_{2R}}| dx \\ &\quad + c \left(\frac{R}{\varrho} \right)^n \left[\frac{|\mu|(B_{2R})}{R^{n-1}} \right]^{\frac{1}{p-1}} + c \left(\frac{R}{\varrho} \right)^n [L_1 \omega(R)]^{\frac{2}{p}} \int_{B_{2R}} (|Du| + s) dx. \end{aligned}$$

Moreover, in the case the vector field $a(\cdot)$ is independent of the variable x , the previous inequality holds without any restriction on R .

Proof. Starting by B_{2R} we define the comparison functions v and w as in (3.33) and (3.23), respectively, and then we compare Du and Dv by mean of (3.37), using (3.6) as basic reference estimate for v , eventually transferred to u :

$$\begin{aligned}
 & \int_{B_\varrho} |Du - (Du)_{B_\varrho}| dx \leq 2 \int_{B_\varrho} |Du - (Dv)_{B_\varrho}| dx \\
 & \leq 2 \int_{B_\varrho} |Dv - (Dv)_{B_\varrho}| dx + 2 \int_{B_\varrho} |Du - Dv| dx \\
 & \leq c \left(\frac{\varrho}{R} \right)^\beta \int_{B_R} |Dv - (Dv)_{B_R}| dx + c \left(\frac{R}{\varrho} \right)^n \int_{B_R} |Du - Dv| dx \\
 & \leq c \left(\frac{\varrho}{R} \right)^\beta \int_{B_R} |Du - (Du)_{B_R}| dx + c \left(\frac{R}{\varrho} \right)^n \int_{B_R} |Du - Dv| dx \\
 (3.42) \quad & \leq 2^{n+1} c \left(\frac{\varrho}{R} \right)^\beta \int_{B_{2R}} |Du - (Du)_{B_{2R}}| dx + c \left(\frac{R}{\varrho} \right)^n \int_{B_R} |Du - Dv| dx.
 \end{aligned}$$

Notice that from the second-last to the last line we estimated as follows:

$$\begin{aligned}
 & \int_{B_R} |Du - (Du)_{B_R}| dx \leq \int_{B_R} |Du - (Du)_{B_{2R}}| dx \\
 (3.43) \quad & + |(Du)_{B_R} - (Du)_{B_{2R}}| \leq 2^{n+1} \int_{B_{2R}} |Du - (Du)_{B_{2R}}| dx.
 \end{aligned}$$

In order to get (3.41) it is now sufficient to estimate the last integral in (3.42) by mean of (3.37) and then to take the radius R_1 such that $L_1 \omega(R_1) \leq 1$. Needless to say, in the case $a(\cdot)$ does not depend on the variable x we can take $\omega(\cdot) \equiv 0$ and therefore no restriction on R is needed. \square

Proof of Theorem 1.1. The proof is divided in three steps. In what follows all the radii considered will be smaller than a certain radius \tilde{R} :

$$(3.44) \quad R \leq \tilde{R}$$

whose final values will be determined towards the end of the proof, i.e. we shall decrease the values of \tilde{R} several times according to our needs, but always in such a way that the resulting determination of \tilde{R} will be still depending only on $n, p, \nu, L, L_1, \omega(\cdot)$. This will finally give the radius \tilde{R} in the statement of Theorem 1.1. From the proof it will be clear that such a restriction is necessary only in the case the vector field $a(\cdot)$ depends on the variable x , while all the inequalities holds with no restriction on R when $a \equiv a(z)$; the final outcome is that estimate (1.6) holds for every ball $B(x_0, 2R) \subseteq \Omega$. This observation also clarifies the last assertion in the statement of Theorem 1.1. Initially, we shall take $\tilde{R} \leq R_1$ where the radius $R_1 \equiv R_1(L_1, \omega(\cdot))$ has been determined in Lemma 3.6 above. In the following we shall consider several constants, in general depending *at least* on the parameters n, p, ν, L ; relevant dependence on additional parameters will be emphasized.

Step 1: Basic dyadic sequence. Referring to estimate (3.41), we select an integer $H \equiv H(n, p, \nu, L) \geq 1$ such that

$$(3.45) \quad c_1 \left(\frac{1}{H} \right)^\beta \leq \frac{1}{4}.$$

We notice that the dependence of H upon n, p, ν, L comes from the similar dependence of β and c_1 presented in Lemma 3.6. Applying (3.41) on arbitrary balls

$B_\varrho \equiv B_{R/2H} \subseteq B_{R/2} \subset B_R$ and using the fact that $\omega(\cdot)$ is non-decreasing we gain

$$(3.46) \quad \begin{aligned} \int_{B_{R/2H}} |Du - (Du)_{B_{R/2H}}| dx &\leq \frac{1}{4} \int_{B_R} |Du - (Du)_{B_R}| dx \\ &+ c \left[\frac{|\mu|(B_R)}{R^{n-1}} \right]^{\frac{1}{p-1}} + c_2 [L_1 \omega(R)]^{\frac{2}{p}} \int_{B_R} (|Du| + s) dx, \end{aligned}$$

where c, c_2 depends only on n, p, ν, L, H and therefore ultimately on n, p, ν, L . By the elementary estimation

$$\int_{B_R} (|Du| + s) dx \leq \int_{B_R} |Du - (Du)_{B_R}| dx + |(Du)_{B_R}| + s$$

estimate (3.46) turns to

$$(3.47) \quad \begin{aligned} \int_{B_{R/2H}} |Du - (Du)_{B_{R/2H}}| dx &\leq \left(\frac{1}{4} + c_2 [L_1 \omega(R)]^{\frac{2}{p}} \right) \int_{B_R} |Du - (Du)_{B_R}| dx \\ &+ c \left[\frac{|\mu|(B_R)}{R^{n-1}} \right]^{\frac{1}{p-1}} + c [L_1 \omega(R)]^{\frac{2}{p}} (|(Du)_{B_R}| + s). \end{aligned}$$

We start reducing the value of \tilde{R} : we take \tilde{R} , depending only on n, p, ν, L, L_1 and $\omega(\cdot)$, small enough in order to get

$$(3.48) \quad c_2 [L_1 \omega(\tilde{R})]^{\frac{2}{p}} \leq \frac{1}{4}.$$

Notice that in order to establish the claimed dependence of \tilde{R} upon the various parameters we have used that c_2 depends in turn on n, p, ν, L . By (3.44), as $\omega(\cdot)$ is non-decreasing, merging (3.47) and (3.48) yields

$$(3.49) \quad \begin{aligned} \int_{B_{R/2H}} |Du - (Du)_{B_{R/2H}}| dx &\leq \frac{1}{2} \int_{B_R} |Du - (Du)_{B_R}| dx \\ &+ c \left[\frac{|\mu|(B_R)}{R^{n-1}} \right]^{\frac{1}{p-1}} + c [L_1 \omega(R)]^{\frac{2}{p}} (|(Du)_{B_R}| + s). \end{aligned}$$

We now fix a ball $B(x_0, 2R) \subseteq \Omega$ as in the statement of Theorem 1.1; for $i \in \{0, 1, 2, \dots\}$, let us define

$$(3.50) \quad B_i := B(x_0, R/(2H)^i) := B(x_0, R_i) \quad \text{and} \quad k_i := |(Du)_{B_i}|.$$

For every integer $m \in \mathbb{N}$ we write

$$(3.51) \quad \begin{aligned} k_{m+1} = \sum_{i=0}^m (k_{i+1} - k_i) + k_0 &\leq \sum_{i=0}^m \int_{B_{i+1}} |Du - (Du)_{B_i}| dx + k_0 \\ &\leq \sum_{i=0}^m (2H)^n \int_{B_i} |Du - (Du)_{B_i}| dx + k_0. \end{aligned}$$

Therefore, defining

$$(3.52) \quad A_i := \int_{B_i} |Du - (Du)_{B_i}| dx$$

inequality (3.51) becomes

$$(3.53) \quad k_{m+1} \leq (2H)^n \sum_{i=0}^m A_i + k_0.$$

To estimate the right hand side of the previous inequality we observe that (3.49) used with $R \equiv R_{i-1}$ yields, whenever $i \geq 1$

$$(3.54) \quad A_i \leq \frac{1}{2}A_{i-1} + c \left[\frac{|\mu|(B_{i-1})}{R_{i-1}^{n-1}} \right]^{\frac{1}{p-1}} + c [L_1 \omega(R_{i-1})]^{\frac{2}{p}} (k_{i-1} + s),$$

with $c \equiv c(n, p, \nu, L)$, where we have taken into account (3.50) and that H depends on n, p, ν, L . We now consider (3.54) for $i \in \{1, \dots, m\}$ and sum up over i , thereby gaining

$$\sum_{i=1}^m A_i \leq \frac{1}{2} \sum_{i=0}^{m-1} A_i + c \sum_{i=0}^{m-1} \left[\frac{|\mu|(B_i)}{R_i^{n-1}} \right]^{\frac{1}{p-1}} + c \sum_{i=0}^{m-1} [L_1 \omega(R_i)]^{\frac{2}{p}} (k_i + s),$$

and therefore

$$\sum_{i=1}^m A_i \leq A_0 + 2c \sum_{i=0}^{m-1} \left[\frac{|\mu|(B_i)}{R_i^{n-1}} \right]^{\frac{1}{p-1}} + 2c \sum_{i=0}^{m-1} [L_1 \omega(R_i)]^{\frac{2}{p}} (k_i + s).$$

Using the last inequality in (3.53) yields, for every integer $m \geq 1$

$$(3.55) \quad k_{m+1} \leq c \left(A_0 + k_0 + \sum_{i=0}^{m-1} \left[\frac{|\mu|(B_i)}{R_i^{n-1}} \right]^{\frac{1}{p-1}} \right) + c \sum_{i=0}^{m-1} [L_1 \omega(R_i)]^{\frac{2}{p}} (k_i + s),$$

as usual for a constant $c \equiv c(n, p, \nu, L)$.

Step 2: Wolff Potentials and Dini-continuity. We notice that

$$\begin{aligned} \sum_{i=0}^{m-1} \left[\frac{|\mu|(B_i)}{R_i^{n-1}} \right]^{\frac{1}{p-1}} &\leq \sum_{i=0}^{\infty} \left[\frac{|\mu|(B_i)}{R_i^{n-1}} \right]^{\frac{1}{p-1}} \\ &\leq \frac{2^{\frac{n-1}{p-1}}}{\log 2} \int_R^{2R} \left[\frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-1}} \right]^{\frac{1}{p-1}} \frac{d\varrho}{\varrho} + \sum_{i=0}^{\infty} \left[\frac{|\mu|(B_{i+1})}{R_{i+1}^{n-1}} \right]^{\frac{1}{p-1}} \\ &\leq \frac{2^{\frac{n-1}{p-1}}}{\log 2} \int_R^{2R} \left[\frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-1}} \right]^{\frac{1}{p-1}} \frac{d\varrho}{\varrho} \\ &\quad + \frac{(2H)^{\frac{n-1}{p-1}}}{\log 2H} \sum_{i=0}^{\infty} \int_{R_{i+1}}^{R_i} \left[\frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-1}} \right]^{\frac{1}{p-1}} \frac{d\varrho}{\varrho} \\ &\leq \left(\frac{2^{\frac{n-1}{p-1}}}{\log 2} + \frac{(2H)^{\frac{n-1}{p-1}}}{\log 2H} \right) \mathbf{W}_{\frac{1}{p}, p}^{\mu}(x_0, 2R). \end{aligned}$$

Recalling the dependence of H determined in (3.45) we conclude that there exists a constant c depending only on n, p, ν, L such that

$$(3.56) \quad \sum_{i=0}^{m-1} \left[\frac{|\mu|(B_i)}{R_i^{n-1}} \right]^{\frac{1}{p-1}} \leq c \mathbf{W}_{\frac{1}{p}, p}^{\mu}(x_0, 2R)$$

holds whenever $m \in \mathbb{N}$. As for the last term in (3.55), using the fact that $\omega(\cdot)$ is non-decreasing, we estimate

$$\begin{aligned} \sum_{i=0}^{m-1} [\omega(R_i)]^{\frac{2}{p}} &\leq \sum_{i=0}^{\infty} [\omega(R_i)]^{\frac{2}{p}} \\ &\leq \frac{1}{\log 2} \int_R^{2R} [\omega(\varrho)]^{\frac{2}{p}} \frac{d\varrho}{\varrho} + \sum_{i=0}^{\infty} [\omega(R_{i+1})]^{\frac{2}{p}} \\ &\leq \frac{1}{\log 2} \int_R^{2R} [\omega(\varrho)]^{\frac{2}{p}} \frac{d\varrho}{\varrho} + \frac{1}{\log 2H} \sum_{i=0}^{\infty} \int_{R_{i+1}}^{R_i} [\omega(\varrho)]^{\frac{2}{p}} \frac{d\varrho}{\varrho} \end{aligned}$$

$$\leq \left(\frac{1}{\log 2} + \frac{1}{\log 2H} \right) \int_0^{2R} [\omega(\varrho)]^{\frac{2}{p}} \frac{d\varrho}{\varrho}.$$

Recalling the definition of $d(\cdot)$ in (1.3), and the fact that $H \geq 1$ we have

$$(3.57) \quad \sum_{i=0}^{m-1} [\omega(R_i)]^{\frac{2}{p}} \leq \frac{2d(2R)}{\log 2}.$$

For later use, we now further restrict the value of \tilde{R} in order to have

$$(3.58) \quad L_1^{\frac{2}{p}} d(2\tilde{R}) \leq 1.$$

In light of inequalities (3.56) and (3.57), (3.55) yields, for every $m \geq 1$

$$(3.59) \quad k_{m+1} \leq c_3 M + c_4 \sum_{i=0}^{m-1} [L_1 \omega(R_i)]^{\frac{2}{p}} k_i,$$

where c_3, c_4 depend only on n, p, ν, L and we have set

$$(3.60) \quad M := \int_{B(x_0, R)} (|Du| + s) dx + \mathbf{W}_{\frac{1}{p}, p}^\mu(x_0, 2R).$$

Notice that, using (3.58) and Hölder's inequality, to get (3.59) from (3.55) we have also estimated

$$(3.61) \quad A_0 + k_0 + d(2R)s \leq c \int_{B(x_0, R)} (|Du| + s) dx.$$

Finally, as in (3.51), by estimating $k_1 = (k_1 - k_0) + k_0 \leq cA_0 + k_0$, and using (3.61), we complement (3.59) with

$$(3.62) \quad k_0 + k_1 \leq c_3 M$$

that we obtain by enlarging c_3 a bit.

Step 3: Induction and conclusion. We restrict the value of \tilde{R} for the last time in order to have

$$(3.63) \quad d(2\tilde{R}) \leq \frac{1}{8c_4 L_1^{2/p}}$$

where $c_4 \equiv c_4(n, p, \nu, L)$ is the constant appearing in (3.59); this, together with the choices (3.48) and (3.63), finally determines \tilde{R} as a positive quantity depending only on $n, p, \nu, L, L_1, \omega(\cdot)$, as required in the statement of the Theorem. We now prove that the following inequality holds whenever $m \geq 1$:

$$(3.64) \quad k_{m+1} \leq 2c_3 M,$$

where c_3 and M have been introduced in (3.59) and (3.60), respectively. We prove (3.64) by induction; the cases $m = -1, 0$ are settled in (3.62); now we assume the validity of (3.64) whenever $m \leq \tilde{m}$, and prove it for $\tilde{m} + 1$. We have, using (3.59), (3.64), (3.57), and (3.63)

$$(3.65) \quad \begin{aligned} k_{\tilde{m}+2} &\leq c_3 M + c_4 \sum_{i=0}^{\tilde{m}} [L_1 \omega(R_i)]^{\frac{2}{p}} k_i \\ &\leq c_3 M + 2c_4 c_3 M \sum_{i=0}^{\tilde{m}} [L_1 \omega(R_i)]^{\frac{2}{p}} \\ &\leq c_3 M + \frac{4c_4 c_3 L_1^{2/p}}{\log 2} d(2R) M \\ &\leq c_3 M + 8c_4 c_3 L_1^{2/p} d(2R) M \\ &\leq c_3 M + c_3 M = 2c_3 M, \end{aligned}$$

so that (3.64) holds whenever $m \in \mathbb{N}$. We now recall that, being Du assumed to be a continuous vector field, for every $x_0 \in \Omega$ it holds that

$$(3.66) \quad |Du(x_0)| = \lim_{m \rightarrow \infty} k_{m+1} \leq 2c_3 M.$$

Merging the last inequality with (3.60) finally yields (1.6). For the case with no x -dependence in the vector field $a(\cdot)$ we just notice that in this case the last term in (3.46) does not appear, and the same happens in (3.49), therefore no smallness assumption of the type (3.48) or (3.63) are required, and all the estimates in the iteration procedure work with no restriction of the type (3.44). The final outcome is that estimate (1.6) holds for every ball $B(x_0, 2R) \subseteq \Omega$. We finally observe that the only point in this proof where we used that $u \in C^1(\Omega)$ was at the very end, in order to assert (3.66) at every point x_0 ; otherwise we could have just assumed $u \in W_{\text{loc}}^{1,p}(\Omega)$ to get (3.66) almost everywhere. \square

4. PARABOLIC ESTIMATES

This section is devoted to the proof of Theorems 1.2 and 1.4, and therefore we shall in general argue under assumptions (1.18) and (1.28), respectively. We observe that assumptions (1.18)₁-(1.18)₂ obviously imply (1.28), and this fact will be implicitly used several times in the following.

4.1. Basic estimates from Nash-Moser's theory. Here we shall emphasize a few properties of solutions $\tilde{u} \in C^0(t_1, t_2; L^2(B(\bar{x}, \gamma))) \cap L^2(t_1, t_2; W^{1,2}(B(\bar{x}, \gamma)))$ for $\gamma > 0$, to homogeneous, non-linear, parabolic equations of the type

$$(4.1) \quad \tilde{u}_t - \operatorname{div} b(x, t, D\tilde{u}) = 0,$$

therefore considered in the basic cylinder $\tilde{Q} \equiv B(\bar{x}, \gamma) \times (t_1, t_2)$. The vector field $b: \tilde{Q} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is supposed to be only Carathéodory regular - in particular the dependence on the coefficients (x, t) is merely measurable. The next result essentially encodes Nash-Moser's regularity for parabolic equations.

Proposition 4.1. *Let $\tilde{u} \in C^0(t_1, t_2; L^2(B(\bar{x}, \gamma))) \cap L^2(t_1, t_2; W^{1,2}(B(\bar{x}, \gamma)))$ be a weak solution to (4.1), under the assumptions (1.28). Then $u \in C_{\text{loc}}^{0,\beta}(\tilde{Q})$ for some $\beta \in (0, 1]$ depending only on n, ν, L , and moreover, there exists a constant c , again depending only on n, ν, L , such that the following inequality:*

$$(4.2) \quad \int_{Q_\varrho} |\tilde{u} - (\tilde{u})_{Q_\varrho}| dx dt \leq c \left(\frac{\varrho}{R} \right)^\beta \int_{Q_R} |\tilde{u} - (\tilde{u})_{Q_R}| dx dt + csR$$

holds whenever $Q_\varrho \subseteq Q_R \subseteq \tilde{Q}$ are parabolic cylinders with the same vertex.

Proof. We shall rely on the precise estimates given by Lieberman in [41, Chapter 6]; this is in turn an extension of the classical Nash-Moser's theory for linear parabolic equations [51, 52]. We first observe that by a simple application of Young's inequality from (1.28) it also follows

$$(4.3) \quad (\nu/2)|z|^2 \leq \langle b(x, t, z), z \rangle + cs^2,$$

where $c \equiv 2L^2/\nu$; compare with (3.35). Therefore applying [41] we have

$$(4.4) \quad \operatorname{osc}_{Q_\varrho} \tilde{u} \leq c \left(\frac{\varrho}{R/2} \right)^\beta \operatorname{osc}_{Q_{R/2}} \tilde{u} + csR \quad \text{for every } \varrho \leq R/2,$$

for constants $c \geq 1$ and $\beta \in (0, 1]$, depending only on n, ν, L . Here we are adopting the standard notation

$$\operatorname{osc}_{Q_\varrho} \tilde{u} = \sup_{Q_\varrho} |\tilde{u}(x, t) - \tilde{u}(\tilde{x}, \tilde{t})| \quad (x, t), (\tilde{x}, \tilde{t}) \in Q_\varrho.$$

Estimate (4.4) maybe necessitates a few clarifications; it can be inferred from [41, Theorem 6.28] once the following things are taken into account. First, one has to adapt the estimates derived in [41, Chapter 6], considering the growth assumption (4.3) and to compare this with [41, (6.20)] and [41, (6.28)]. Second, the estimates in [41, Chapter 6] are often stated for linear operators, but the linearity is actually irrelevant, as growth conditions are the only relevant thing, therefore all the estimates can be derived for general quasilinear operators; see the comments at the beginning of [41, Section 6.5]. Again from [41, Theorem 6.17] we gain

$$(4.5) \quad \sup_{Q_{R/2}} |\tilde{u}| \leq c \int_{Q_R} |\tilde{u}| dx dt + csR,$$

again for a constant c depending only on n, ν, L , and whenever $Q_R \subseteq \tilde{Q}$. With $m \in \mathbb{R}$ being a fixed number, we observe that $\tilde{u} - m$ still solves (4.1), therefore estimate (4.5) gives

$$(4.6) \quad \sup_{Q_{R/2}} |\tilde{u} - m| \leq c \int_{Q_R} |\tilde{u} - m| dx dt + csR.$$

On the other hand we have

$$(4.7) \quad \text{osc}_{Q_{R/2}}(\tilde{u} - m) \leq 2 \sup_{Q_{R/2}} |\tilde{u} - m| \leq c \int_{Q_R} |\tilde{u} - m| dx dt + csR,$$

and, in turn applying estimate (4.4) to $\tilde{u} - m$ we find, by mean of (4.7)

$$\text{osc}_{Q_e} \tilde{u} = \text{osc}_{Q_e}(\tilde{u} - m) \leq c \left(\frac{\rho}{R} \right)^\beta \int_{Q_R} |\tilde{u} - m| dx dt + csR.$$

Noticing that

$$\int_{Q_e} |\tilde{u} - (\tilde{u})_{Q_e}| dx dt \leq \text{osc}_{Q_e} \tilde{u}$$

we finally obtain

$$\int_{Q_e} |\tilde{u} - (\tilde{u})_{Q_e}| dx dt \leq c \left(\frac{\rho}{R} \right)^\beta \int_{Q_R} |\tilde{u} - m| dx dt + csR.$$

At this stage (4.2) follows choosing $m = (\tilde{u})_{Q_R}$. \square

4.2. Comparison estimates. In the rest of Section 4.2 we keep fixed a symmetric parabolic cylinder $Q(x_0, t_0, 2R) \equiv Q_{2R} \subset \Omega_T$. We start considering the unique solution

$$(4.8) \quad w \in C^0(t_0 - 4R^2, t_0; L^2(B(x_0, 2R))) \cap L^2(t_0 - 4R^2, t_0; W^{1,2}(B(x_0, 2R)))$$

to the following Cauchy-Dirichlet problem:

$$(4.9) \quad \begin{cases} w_t - \text{div } a(x, t, Dw) = 0 & \text{in } Q_{2R} \\ w = u & \text{on } \partial_{\text{par}} Q_{2R}. \end{cases}$$

Lemma 4.1. *Let $u \in C^0(-T, 0; L^2(\Omega)) \cap L^2(-T, 0; W^{1,2}(\Omega))$ be a solution to (1.17) with $\mu \in L^2(\Omega_T)$, under the assumptions (1.28), and let w be defined in (4.8)-(4.9). Then there exists a constant $c = c(n, \nu)$ such that the following estimates hold:*

$$(4.10) \quad \int_{Q_{2R}} |Du - Dw| dx dt \leq \frac{c|\mu|(Q_{2R})}{R^{N-1}},$$

$$(4.11) \quad \int_{Q_{2R}} |u - w| dx dt \leq \frac{c|\mu|(Q_{2R})}{R^{N-2}}.$$

Proof. Step 1: Universal estimate. Here we assume $(x_0, t_0) = (0, 0)$ and $2R = 1$ - that is $Q_{2R} \equiv Q_1$ - and that $|\mu|(Q_1) = 1$, and prove that the following universal inequality holds:

$$(4.12) \quad \int_Q |Du - Dw| dx dt \leq \frac{c(n)}{\nu}.$$

Our starting point here will be the parabolic estimates developed in [9]. As described in Section 2, we use the Steklov-averages formulation of both (1.17) and (4.9)₁, i.e., for every $t \in (-1, 0)$ and $t + h < 0$ there holds

$$(4.13) \quad \int_B \left(\partial_t u_h(\cdot, t) \varphi + \langle [a(\cdot, t, Du)]_h, D\varphi \rangle \right) dx = \int_B \varphi \mu_h(x, t) dx$$

and

$$(4.14) \quad \int_B \left(\partial_t w_h(\cdot, t) \varphi + \langle [a(\cdot, t, Dw)]_h, D\varphi \rangle \right) dx = 0$$

for every $\varphi \in C_0^\infty(B)$ with compact support, and by density whenever $\varphi \in W_0^{1,2}(B)$. The initial datum of w_h is here taken in the sense of $L^2(B)$, which means that $w_h(\cdot, -1) \rightarrow u(\cdot, -1)$ in $L^2(B)$ when $h \downarrow 0$, and since $u \in C^0([-1, 0]; L^2(B))$, this implies in particular that

$$(4.15) \quad \lim_{h \downarrow 0} \|(u_h - w_h)(\cdot, -1)\|_{L^2(B)} = 0.$$

With the notation fixed in (3.27), by $\Psi_k: \mathbb{R} \rightarrow \mathbb{R}$ we denote the following primitive $\Psi_k(s) := \int_0^s \Phi_k(\kappa) d\kappa$ for $\tau \in \mathbb{R}$, that for later use we compute explicitly:

$$\Psi_k(\tau) = \begin{cases} \frac{1}{2} + (\tau - k - 1) & \tau \geq k + 1 \\ \frac{1}{2}(\tau - k)^2 & k < \tau < k + 1 \\ 0 & -k \leq \tau \leq k \\ \frac{1}{2}(\tau + k)^2 & -k - 1 < \tau < -k \\ \frac{1}{2} - (\tau + k + 1) & \tau \leq -k - 1, \end{cases}$$

and note that $\Psi_k \geq 0$. By testing the difference Steklov-formulations (4.13), (4.14)

$$(4.16) \quad \begin{aligned} & \int_B \left(\partial_t (u_h - w_h)(x, t) \varphi + \langle [a(\cdot, t, Du)]_h - [a(\cdot, t, Dw)]_h, D\varphi \rangle \right) dx \\ &= \int_B \varphi \mu_h(x, t) dx, \end{aligned}$$

with the choice $\varphi(x, t) := \zeta(t) \Phi_k(u_h - w_h)(x, t)$, $x \in B$, where $\zeta(\cdot)$ is a smooth function, and then integrating the resulting equality on $(-1, 0)$ with respect to t we obtain

$$(4.17) \quad \begin{aligned} & \int_Q \partial_t (u_h - w_h) \Phi_k(u_h - w_h) \zeta dx dt \\ &+ \int_Q \langle [a(\cdot, Du)]_h - [a(\cdot, Dw)]_h, D\Phi_k(u_h - w_h) \rangle \zeta dx dt \\ &= \int_Q \Phi_k(u_h - w_h) \zeta \mu_h dx dt. \end{aligned}$$

With $\tau \in (-1, 0)$ now choose $\zeta \in W^{1,\infty}(\mathbb{R})$ as follows:

$$\zeta_\varepsilon(\kappa) \equiv \zeta(\kappa) = \begin{cases} 1 & \kappa \leq \tau \\ 1 - \frac{1}{\varepsilon}(t - \kappa) & \tau < \kappa < \tau + \varepsilon \\ 0 & \kappa > \tau + \varepsilon. \end{cases}$$

Then, the first integral in (4.17) can be rewritten in the form

$$\begin{aligned}
 & \int_Q \partial_t(u_h - w_h) \Phi_k(u_h - w_h) \zeta \, dx \, dt \\
 &= \int_Q \partial_t [\Psi_k(u_h - w_h) \zeta] \, dx \, dt - \int_Q \Psi_k(u_h - w_h) \zeta_t \, dx \, dt \\
 (4.18) \quad &= - \int_B \Psi_k(u_h - w_h)(x, -1) \, dx - \int_Q \Psi_k(u_h - w_h) \zeta_t \, dx \, dt.
 \end{aligned}$$

Taking into account the special choice of ζ we see that the second integral appearing on the right-hand side of (4.18) converges to $\int_B \Psi_k(u_h - w_h)(x, \tau) \, dx$ when $\varepsilon \downarrow 0$ for almost every $\tau \in (-1, 0)$, while the first integral vanishes as $h \downarrow 0$ in light of (4.15). As we have seen above the initial datum is taken in the sense of $L^2(B)$. Combining (4.18) with (4.17) and letting first $\varepsilon \downarrow 0$ and then $h \downarrow 0$ in (4.18) we obtain

$$\begin{aligned}
 & \int_B \Psi_k(u - w)(x, \tau) \, dx \\
 &+ \int_{-1}^{\tau} \int_B \langle a(x, t, Du) - a(x, t, Dw), D\Phi_k(u - w) \rangle \, dx \, dt = \int_{-1}^{\tau} \int_B \Phi_k(u - w) \mu \, dx \, dt
 \end{aligned}$$

for almost every every choice of $\tau \in (-1, 0)$. This leads to

$$\begin{aligned}
 & \sup_{-1 < \tau < 0} \int_B \Psi_k(u - w)(x, \tau) \, dx \\
 &+ \int_Q \langle a(x, t, Du) - a(x, t, Dw), D\Phi_k(u - w) \rangle \, dx \, dt \\
 (4.19) \quad & \leq \int_Q |\Phi_k(u - w)| |\mu| \, dx \, dt.
 \end{aligned}$$

We first exploit the properties of Ψ_k in order to estimate the first term of the left-hand side from below. We start with the obvious estimate

$$(4.20) \quad \sup_{-1 < \tau < 0} \int_B \Psi_0(u - w)(\cdot, \tau) \, dx \leq |\mu|(Q) \leq 1,$$

which follows by applying (4.19) for $k = 0$; note also that $|\mu|(Q) = 1$ by assumption. With τ being fixed, decomposing

$$B = \{x \in B : |(u - w)(x, \tau)| \leq 1\} \cup \{x \in B : |(u - w)(x, \tau)| > 1\} =: B'_\tau \cup B''_\tau$$

we have for almost every $\tau \in (-1, 0)$ that

$$\begin{aligned}
 & \int_B \Psi_0(u - w)(x, \tau) \, dx \\
 (4.21) \quad &= \int_{B'_\tau} \frac{1}{2} |(u - w)(x, \tau)|^2 \, dx + \int_{B''_\tau} |(u - w)(x, \tau)| \, dx - \frac{1}{2} |B''_\tau|
 \end{aligned}$$

holds. On the other hand we trivially have $\int_{B'_\tau} |(u - w)(x, \tau)| \, dx \leq |B|$, so that merging the last estimate with (4.20)-(4.21) we have

$$\int_B |(u - w)(x, \tau)| \, dx \leq \frac{1}{2} |B| + |\mu|(Q) \leq c(n)$$

which holds again for almost every $\tau \in (-1, 0)$, and therefore we conclude with the $L^\infty - L^1$ -bound

$$(4.22) \quad \|u - w\|_{L^\infty(-1, 0; L^1(B))} := \sup_{-1 < \tau < 0} \int_B |(u - w)(x, \tau)| \, dx \leq c(n).$$

We now start exploiting the second integral appearing on the left-hand side of (4.19). For this we define

$$A_k := \{(x, t) \in Q : k \leq |u(x, t) - w(x, t)| < k + 1\}.$$

Then, using the fact that $D\Phi_k(u - w) = Du - Dw$ on A_k and $D\Phi_k(u - w) = 0$ otherwise together with (1.28)₂, $\Phi_k(\cdot) \leq 1$ and $|\mu|(Q) = 1$ in (4.19) we obtain

$$\nu \int_{A_k} |Du - Dw|^2 dx dt \leq \int_{A_k} \langle a(x, t, Du) - a(x, t, Dw), D(u - w) \rangle dx dt \leq 1$$

so that the estimate

$$(4.23) \quad \int_{A_k} |Du - Dw|^2 dx dt \leq \frac{1}{\nu}$$

holds for every $k \in \{0, 1, 2, \dots\}$. Now, for $\lambda > 1$ we estimate

$$(4.24) \quad \begin{aligned} \int_Q \frac{|Du - Dw|^2}{(1 + |u - w|)^\lambda} dx dt &= \sum_{k=0}^{\infty} \int_{A_k} \frac{|Du - Dw|^2}{(1 + |u - w|)^\lambda} dx dt \\ &\leq \sum_{k=0}^{\infty} \frac{1}{(1 + k)^\lambda} \int_{A_k} |Du - Dw|^2 dx dt \leq \frac{\lambda}{\nu(\lambda - 1)}. \end{aligned}$$

Here we have used in the last line the uniform estimate (4.23). This allows us to argue as follows: For almost every $\tau \in (-1, 0)$ we have by the Cauchy-Schwartz inequality that

$$\begin{aligned} \int_B |D(u - w)(x, \tau)| dx &= \int_B \frac{|D(u - w)(x, \tau)|}{(1 + |(u - w)(x, \tau)|)^{\lambda/2}} (1 + |(u - w)(x, \tau)|)^{\lambda/2} dx \\ &\leq \left(\int_B \frac{|D(u - w)(x, \tau)|^2}{(1 + |(u - w)(x, \tau)|)^\lambda} dx \right)^{\frac{1}{2}} \left(\int_B (c(n) + |(u - w)(x, \tau)|)^\lambda dx \right)^{\frac{1}{2}}. \end{aligned}$$

Integrating the preceding inequality with respect to τ on $(-1, 0)$ and using Cauchy-Schwartz with respect to τ together with (4.24) then yields

$$(4.25) \quad \begin{aligned} &\int_Q |Du - Dw| dx dt \\ &\leq \int_{-1}^0 \left(\int_B \frac{|D(u - w)(x, t)|^2}{(1 + |(u - w)(x, t)|)^\lambda} dx \right)^{\frac{1}{2}} \left(\int_B (1 + |(u - w)(x, t)|)^\lambda dx \right)^{\frac{1}{2}} dt \\ &\leq \left(\int_Q \frac{|D(u - w)|^2}{(1 + |u - w|)^\lambda} dx dt \right)^{\frac{1}{2}} \left(\int_{-1}^0 \int_B (1 + |(u - w)(x, t)|)^\lambda dx dt \right)^{\frac{1}{2}} \\ &\leq \sqrt{\frac{2^{\lambda-1}\lambda}{\nu(\lambda-1)}} \left[1 + \left(\int_{-1}^0 \|(u - w)(x, t)\|_{L^\lambda(B)}^\lambda dt \right)^{\frac{1}{2}} \right]. \end{aligned}$$

At this stage we use a well-known version of the Gagliardo-Nirenberg embedding theorem which reads in our setting as follows:

$$\|(u - w)(x, t)\|_{L^\lambda(B)} \leq c(n) \|D(u - w)(x, t)\|_{L^1(B)}^\theta \|(u - w)(x, t)\|_{L^1(B)}^{1-\theta},$$

which holds for every choice of θ and λ satisfying

$$0 \leq \theta \leq 1, \quad 1 < \lambda < \infty, \quad \frac{1}{\lambda} = 1 - \frac{\theta}{n}.$$

We note that we have $\lambda > 1$ at our disposal to ensure $\theta \in (0, 1)$. Recalling the $L^\infty - L^1$ -estimate (4.22) the second integral on the right-hand side of the preceding inequality is bounded by $c(n)^{1-\theta}$, and therefore

$$\int_{-1}^0 \|(u - w)(x, t)\|_{L^\lambda(B)}^\lambda dt \leq c(n)^{1+\lambda(1-\theta)} \int_{-1}^0 \|D(u - w)(x, t)\|_{L^1(B)}^{\lambda\theta} dt.$$

We now perform the choice of θ and λ . In order to have $\theta\lambda = 1$, i.e. $\theta = \frac{1}{\lambda}$, the identity relating θ and λ yields $\lambda = \frac{n+1}{n}$, so that $\theta = \frac{n}{n+1}$. This implies that

$$\int_{-1}^0 \|(u-w)(x, t)\|_{L^{1+\frac{1}{n}}(B)}^{1+\frac{1}{n}} dt \leq c(n) \int_Q |D(u-w)| dx dt.$$

Inserting this last estimate in (4.25), and using Young's inequality we arrive at the following inequality, which in turn implies (4.12)

$$\begin{aligned} \int_Q |Du - Dw| dx dt &\leq \frac{c(n)}{\sqrt{\nu}} \left[1 + \left(\int_Q |Du - Dw| dx dt \right)^{\frac{1}{2}} \right] \\ &\leq \frac{c(n)}{\nu} + \frac{1}{2} \int_Q |Du - Dw| dx dt. \end{aligned}$$

Step 2: General case and scaling. Similarly to the elliptic case we first reduce to the case $Q_{2R} = Q(x_0, t_0; 2R) \equiv Q$ by changing variables and passing to new solutions and vector fields. More precisely, for $(\tilde{x}, \tilde{t}) \in Q$

$$(4.26) \quad \begin{cases} \tilde{u}(\tilde{x}, \tilde{t}) := \frac{u(x_0 + 2R\tilde{x}, t_0 + 4R^2\tilde{t})}{2R} \\ \tilde{w}(\tilde{x}, \tilde{t}) := \frac{w(x_0 + 2R\tilde{x}, t_0 + 4R^2\tilde{t})}{2R} \\ \tilde{a}(\tilde{x}, \tilde{t}, z) := a(x_0 + 2R\tilde{x}, t_0 + 4R^2\tilde{t}, z) \\ \tilde{\mu}(\tilde{x}, \tilde{t}) := 2R\mu(x_0 + 2R\tilde{x}, t_0 + 4R^2\tilde{t}). \end{cases}$$

Then it follows that $\tilde{u}_{\tilde{t}} - \operatorname{div} \tilde{a}(\tilde{x}, \tilde{t}, D\tilde{u}) = \tilde{\mu}$ and $\tilde{w}_{\tilde{t}} - \operatorname{div} \tilde{a}(\tilde{x}, \tilde{t}, D\tilde{w}) = 0$ in Q , and $\tilde{u} = \tilde{w}$ on $\partial_{\text{par}}Q$. Furthermore the new vector field $\tilde{a}(\cdot)$ satisfies (1.28). To remove the additional assumption $|\mu|(Q) = 1$ we assume, as we can, that $A := |\mu|(Q) > 0$ - otherwise there is nothing to prove since the strict monotonicity of the vector field would imply $u \equiv w$ - and then we re-scale $u, w, a(\cdot)$ and μ as follows: $\bar{u} := A^{-1}u$, $\bar{w} := A^{-1}w$, $\bar{a}(x, t, z) := A^{-1}a(x, t, Az)$, $\bar{\mu} := A^{-1}\mu$. Then, it is easily verified that \bar{u} is a solution of $\bar{u}_{\tilde{t}} - \operatorname{div} \bar{a}(x, t, \bar{u}) = \bar{\mu}$ on Q , that \bar{w} solves $\bar{w}_{\tilde{t}} - \operatorname{div} \bar{a}(x, t, \bar{w}) = 0$ on Q , that $\bar{u} = \bar{w}$ on $\partial_{\text{par}}Q$ and $|\bar{\mu}|(Q) = 1$. Moreover, $\bar{a}(\cdot)$ fulfills the strict monotonicity condition (1.28) which is the only one used in a quantitative way in the derivation of the universal comparison estimate (4.12). Therefore estimate (4.12) applied to $\bar{u} - \bar{w}$ yields

$$\nu \int_Q |D\bar{u} - D\bar{w}| dx dt \leq c(n),$$

and re-scaling back this inequality from $\bar{u} - \bar{w}$ to $u - w$ leads us to the comparison estimate (4.10) with $R = 1$, and by the previous scaling (4.10) is proved in the general case. In order to get (4.11) it is sufficient to recall that $(u - w)(\cdot, t) \in W_0^{1,2}(B(x_0, 2R))$ for almost every $t \in (t_0 - 4R^2, t_0)$ and therefore applying Poincaré's inequality slicewise and integrating we gain

$$\int_{Q(x_0, t_0; 2R)} |u - w| dx dt \leq cR \int_{Q(x_0, t_0; 2R)} |Du - Dw| dx dt$$

which together with (4.10) implies (4.11). This completes the proof. \square

We now proceed with a further comparison estimate; after having defined the comparison solution w in (4.8) we define the unique solution

$$(4.27) \quad v \in C^0(t_0 - R^2, t_0; L^2(B(x_0, R))) \cap L^2(t_0 - R^2, t_0; W^{1,2}(B(x_0, R)))$$

of the following Cauchy-Dirichlet problem:

$$(4.28) \quad \begin{cases} v_t - \operatorname{div} a(x_0, t, Dv) = 0 & \text{in } Q_R \\ v = w & \text{on } \partial_{\text{par}} Q_R. \end{cases}$$

We recall the reader's attention on the fact that in (4.28) we have frozen the coefficients only with respect to the space variable.

Lemma 4.2. *Let $u \in C^0(-T, 0; L^2(\Omega)) \cap L^2(-T, 0; W^{1,2}(\Omega))$ be a solution to (1.17) with $\mu \in L^1(\Omega_T)$, under the assumptions (1.28), let w be defined in (4.8)-(4.9), and finally let v be defined in (4.27)-(4.28). Then there exists a constant $c \equiv c(n, \nu, L)$ such that the following estimate holds:*

$$(4.29) \quad \begin{aligned} \int_{Q_R} |Du - Dv| dx dt &\leq c[1 + L_1\omega(R)] \frac{|\mu|(Q_{2R})}{R^{N-1}} \\ &\quad + cL_1\omega(R) \int_{Q_{2R}} (|Du| + s) dx dt. \end{aligned}$$

Proof. Let us first get the estimate

$$(4.30) \quad \int_{Q_R} |Dv - Dw|^2 dx dt \leq cL_1^2[\omega(R)]^2 \int_{Q_R} (|Dw| + s)^2 dx dt.$$

The proof consists of a rather standard comparison argument which we report here for the sake of completeness. The following computations are formal, as they would need the existence of the time derivatives for both v and w ; on the other hand they can be made rigorous by using the Steklov-averages formulation exactly as done for (4.17)-(4.18). Using the fact that both v and w are solutions yields

$$(4.31) \quad \begin{aligned} (v - w)_t - \operatorname{div} (a(x_0, t, Dv) - a(x_0, t, Dw)) \\ = \operatorname{div} (a(x_0, t, Dw) - a(x, t, Dw)). \end{aligned}$$

By testing the weak form of the previous equality with $v - w$ - here we need to pass to Steklov averages - we get, after standard parabolic manipulations and the fact that v and w agree on $\partial_{\text{par}} Q_R$, that

$$\begin{aligned} &\sup_{t_0 - R^2 < t < t_0} \int_{B(x_0, R)} |v - w|^2(x, t) dx \\ &\quad + \int_{Q_R} \langle a(x_0, t, Dv) - a(x_0, t, Dw), Dv - Dw \rangle dx dt \\ &\leq \left| \int_{Q_R} \langle a(x_0, t, Dw) - a(x, t, Dw), Dv - Dw \rangle dx dt \right|. \end{aligned}$$

Discarding the first term in the previous inequality, using (1.28) and (1.18)₃, yields

$$\int_{Q_R} |Dv - Dw|^2 dx dt \leq cL_1\omega(R) \int_{Q_R} (|Dw| + s)|Dv - Dw| dx dt.$$

At this stage (4.30) follows by the last estimate and a standard use of Young's inequality. The next step is to recall a higher integrability estimate which follows from the application of Gehring's lemma in the parabolic setting. The statement can be found in several papers see for instance [53], once taking into account the growth conditions (1.18) considered here. There exist constants $c \geq 1$ and $\chi > 2$, depending only on n, ν, L , such that the following inequality:

$$\left(\int_{Q_\varrho} |Dw|^\chi dx dt \right)^{\frac{1}{\chi}} \leq c \left(\int_{Q_{2\varrho}} (|Dw| + s)^2 dx dt \right)^{\frac{1}{2}}$$

holds whenever $Q_{2\rho} \subset Q_{2R}$. We are therefore in position to apply Lemma 3.1, which allows to establish

$$(4.32) \quad \left(\int_{Q_R} |Dw|^2 dx dt \right)^{\frac{1}{2}} \leq c \int_{Q_{2R}} (|Dw| + s) dx dt.$$

Combining (4.30) and (4.32) via Hölder's inequality we gain

$$\int_{Q_R} |Dv - Dw| dx dt \leq L_1 \omega(R) \int_{Q_{2R}} (|Dw| + s) dx dt.$$

Again combining the last estimate with (4.10), and actually using it twice, finally gives (4.29). \square

We conclude this section with yet another a priori estimate, in which, *for the first and only time* in this section we use the complete assumptions (1.18) rather than the weaker ones in (1.28).

Proposition 4.2. *Let v be defined in (4.27)-(4.28); there exist constants $c \geq 1$ and $\beta \in (0, 1]$, depending only on n, ν, L , such that the following inequality:*

$$(4.33) \quad \int_{Q_\rho} |Dv - (Dv)_{Q_\rho}| dx dt \leq c \left(\frac{\rho}{R} \right)^\beta \int_{Q_R} |Dv - (Dv)_{Q_R}| dx dt$$

holds whenever $Q_\rho \subseteq Q_R$ is a parabolic cylinder with the same vertex of Q_R . More in general, the inequality

$$(4.34) \quad \int_{Q_\rho} |D_\xi v - (D_\xi v)_{Q_\rho}| dx dt \leq c \left(\frac{\rho}{R} \right)^\beta \int_{Q_R} |D_\xi v - (D_\xi v)_{Q_R}| dx dt$$

holds whenever $\xi \in \{1, \dots, n\}$ and $\rho \leq R$.

Proof. We recall the difference quotients argument developed for instance in [17, Chapter 8]. We indeed have that v is higher differentiable in a smaller cylinder: $v \in L^2(t_0 - R^2/4, t_0; W^{2,2}(B(x_0, R/2)))$ and moreover, whenever $\xi \in \{1, \dots, n\}$ we have $D_\xi v \in C^0(t_0 - R^2/4, t_0; L^2(B(x_0, R/2)))$. Finally, the function $D_\xi v$ solves the following differentiated equation:

$$(D_\xi v)_t - \operatorname{div}(\tilde{a}(x, t) D D_\xi v) = 0,$$

weakly in $Q(x_0, t_0; R/2) \equiv Q_{R/2}$, where the matrix $\tilde{a}_{ij}(x, t) := D_{z_j} a_i(x_0, t, Dv(x, t))$ has measurable entries. By assumptions (1.18) we have that the following monotonicity and growth conditions are satisfied:

$$\nu |z|^2 \leq \langle \tilde{a}(x, t) z, z \rangle, \quad |\tilde{a}(x, t) z| \leq L |z|$$

whenever $z \in \mathbb{R}^n$, and $(x, t) \in Q_{R/2}$. Therefore, if we let $b(x, t, z) := \tilde{a}(x, t) z$, we may apply Proposition 4.1 to $\tilde{u} \equiv D_\xi v$ in the cylinder $Q_{R/2}$, thereby obtaining

$$(4.35) \quad \begin{aligned} \int_{Q_\rho} |D_\xi v - (D_\xi v)_{Q_\rho}| dx dt &\leq c \left(\frac{\rho}{R/2} \right)^\beta \int_{Q_{R/2}} |D_\xi v - (D_\xi v)_{Q_{R/2}}| dx dt \\ &\leq c \left(\frac{\rho}{R} \right)^\beta \int_{Q_R} |D_\xi v - (D_\xi v)_{Q_R}| dx dt \end{aligned}$$

that holds whenever $\rho \leq R/2$, and where again $c \equiv c(n, \nu, L)$; we notice that the last estimate in (4.35) has been obtained arguing as for (3.43). Finally, the same inequality follows for $\rho \in (R/2, R]$, trivially, and therefore (4.34) is completely established. In turn, since $\xi \in \{1, \dots, n\}$ is arbitrary (4.34) implies (4.33) and the proof is complete. \square

4.3. Proof of Theorems 1.2-1.3. We start by Theorem 1.2. In Sections 4.1-4.2 we have built the necessary set up in order to adapt the elliptic proof of Section 3.4 to the parabolic case, and we shall therefore outline the relevant modifications. Combining estimates (4.29) and (4.33) as done in Lemma 3.6 with estimates (3.6) and (3.37), we have the analogue of estimate (3.41)

$$(4.36) \quad \begin{aligned} \int_{Q_\varrho} |Du - (Du)_{Q_\varrho}| dx dt &\leq c_1 \left(\frac{\varrho}{R}\right)^\beta \int_{Q_{2R}} |Du - (Du)_{Q_{2R}}| dx dt \\ &+ c \left(\frac{R}{\varrho}\right)^N \frac{|\mu|(Q_{2R})}{R^{N-1}} + c_2 \left(\frac{R}{\varrho}\right)^N [L_1\omega(R)] \int_{Q_{2R}} (|Du| + s) dx dt, \end{aligned}$$

where the constants $c, c_1, c_2 \geq 1$ and $\beta \in (0, 1]$ depend only on n, ν, L ; the last inequality holds whenever $Q_\varrho \subseteq Q_R$ are backward parabolic cylinders sharing the same vertex, and with $R < R_1$ where $R_1 \equiv R_1(L_1, \omega(\cdot)) > 0$ being a suitable small radius; the restriction on R is not necessary when $a(\cdot)$ is independent of x , although $a(\cdot)$ can be still depending on the time variable t . By (4.36), choosing $H \in \mathbb{N}$ - similarly to (3.45) - large enough to have $c_1/H^\beta \leq 1/4$, and taking $R \leq \tilde{R} \leq R_1$ such that $c_2(2H)^N [L_1\omega(R)] \leq 1/4$ similarly to (3.48), and proceeding as for (3.46)-(3.49) we obtain

$$(4.37) \quad \begin{aligned} \int_{Q_{R/H}} |Du - (Du)_{Q_{R/H}}| dx dt &\leq \frac{1}{2} \int_{Q_{2R}} |Du - (Du)_{Q_{2R}}| dx dt \\ &+ \frac{c|\mu|(Q_{2R})}{R^{N-1}} + cL_1\omega(R) [|Du|_{Q_{2R}} + s]. \end{aligned}$$

The last inequality is again valid for cylinders with same vertex and for a constant c depending on n, ν, L , with the restriction $R \leq \tilde{R}$. We now just have to follow the scheme after (3.49). Specifically, we let

$$Q_i := Q(x_0, t_0; R/(2H)^i), \quad k_i := |(Du)_{Q_i}|;$$

letting - as in (3.52) -

$$A_i := \int_{Q_i} |Du - (Du)_{Q_i}| dx dt$$

we apply (4.37) with $Q_R \equiv Q_i$, and iterating as after (3.53), we finally achieve estimate (1.23).

As for Theorem 1.3, to obtain (1.25) we argue as follows. We start by (4.34); this holds directly for w defined in (4.8)-(4.9), recall that in this case $v \equiv w$ as we do not have x -dependence and we do not have to freeze at x_0 . Next we combine (4.34) directly with (4.10) and obtain the following analogue of (4.36):

$$(4.38) \quad \begin{aligned} \int_{Q_\varrho} |D_\xi u - (D_\xi u)_{Q_\varrho}| dx dt &\leq c_1 \left(\frac{\varrho}{R}\right)^\beta \int_{Q_{2R}} |D_\xi u - (D_\xi u)_{Q_{2R}}| dx dt \\ &+ c \left(\frac{R}{\varrho}\right)^N \frac{|\mu|(Q_{2R})}{R^{N-1}}. \end{aligned}$$

Now (1.25) follows by iterating as in the proof of Theorem 1.2, by setting this time

$$Q_i := Q(x_0, t_0; R/(2H)^i), \quad k_i := |(D_\xi u)_{Q_i}|$$

and

$$A_i := \int_{Q_i} |D_\xi u - (D_\xi u)_{Q_i}| dx dt.$$

Worth remarking differences are that in (3.60) the value of it is now

$$M := \int_{Q(x_0, t_0; R)} (|D_\xi u| + s) dx dt + \mathbf{I}_1^\mu(x_0, t_0; 2R)$$

and this comes from the fact that, being now $\omega(\cdot)$ we also have $d(2R) = 0$ and therefore (3.61) can be replaced by

$$A_0 + k_0 \leq c \int_{Q(x_0, t_0; R)} (|D_\xi u| + s) dx dt.$$

4.4. Proof of Theorem 1.4. The main difference with respect to the proof of Theorem 1.2 is that we just need to use estimate (4.11), and no other intermediate comparison inequality of the type (4.29). This is essentially the reason why Theorem 1.4 applies - as expected for zero-order estimates - to operators with measurable coefficients. As for the proof, we first obtain the estimate

$$(4.39) \quad \begin{aligned} \int_{Q_\varrho} |u - (u)_{Q_\varrho}| dx dt &\leq c_1 \left(\frac{\varrho}{R}\right)^\beta \int_{Q_{2R}} |u - (u)_{Q_{2R}}| dx dt \\ &\quad + c \left(\frac{R}{\varrho}\right)^N \frac{|\mu|(Q_{2R})}{R^{N-2}} + cRs \end{aligned}$$

which is the analogue of (4.38) and of (4.37) when no space coefficients come into the play. As usual, (4.39) works for backward cylinders with the same vertex, and the constants $c, c_1 \geq 1$ and $\beta \in (0, 1]$ depending only on n, ν, L . Estimate (4.39) can be derived by first applying Proposition 4.1 with $\tilde{u} \equiv w$ and $b(\cdot) \equiv a(\cdot)$, thereby getting the reference estimate

$$(4.40) \quad \int_{Q_\varrho} |w - (w)_{Q_\varrho}| dx dt \leq c \left(\frac{\varrho}{R}\right)^{N+\beta} \int_{Q_{2R}} |w - (w)_{Q_{2R}}| dx dt + cRs,$$

whenever $Q_\varrho \subseteq Q_{2R}$ is a cylinder with the vertex of Q_{2R} , and where $c \geq 1$ and $\beta \in (0, 1]$ depend only on n, ν, L . Combining this last estimate with (4.11) in the same way of Lemma 3.6 - and exactly for (4.36) - finally yields (4.39). By (4.39), choosing H to have $c_1/H^\beta \leq 1/2$, we obtain, again for $c \equiv c(n, \nu, L)$

$$(4.41) \quad \int_{Q_{R/H}} |u - (u)_{Q_{R/H}}| dx dt \leq \frac{1}{2} \int_{Q_{2R}} |u - (u)_{Q_{2R}}| dx dt + \frac{c|\mu|(Q_{2R})}{R^{N-2}} + cRs.$$

The last estimate can be iterated as done for (4.37), as after (3.49); we let

$$Q_i := Q(x_0, t_0; R/(2H)^i), \quad k_i := |(u)_{Q_i}|$$

and finally

$$A_i := \int_{Q_i} |u - (u)_{Q_i}| dx dt$$

and apply (4.41) with $Q_R \equiv Q_i$, iterating as after (3.53), thereby concluding the proof.

Remark 4.1 (Zero order elliptic estimate). The method of proof used for Theorem 1.4 easily leads to an alternative proof of the zero order estimate (1.14) - even with $\gamma = 1$ - from [35, 56] for the elliptic case (1.1) under the assumptions

$$(4.42) \quad \begin{cases} |a(x, z)| \leq L(|z| + s)^{p-1} \\ \nu|z_2 - z_1|^p \leq \langle a(x, z_2) - a(x, z_1), z_2 - z_1 \rangle \end{cases}$$

whenever $z_1, z_2 \in \mathbb{R}^n$ and $x \in \Omega$; compare with (1.28). We observe that this proof works also in the case of general signed measures, and the final outcome is

$$(4.43) \quad |u(x_0)| \leq c \int_{B(x_0, R)} (|u| + Rs) dx + c\mathbf{W}_{1,p}^\mu(x_0, 2R).$$

The proof in question seems to be slightly easier than those proposed up to now, although the assumptions (4.42) are slightly stronger than those in [56]. For the proof, following the scheme of the proof of Theorem 1.4, we essentially need to have two ingredients: a reference decay estimate for comparison function w introduced

in (3.23), and a suitable comparison estimate between u and w . This last one has already been proved in (3.25), while following for instance [48, Theorem 5.1], and using as a starting point the proof of the decay estimate in [24, Theorem 7.7], we have that the following estimate holds whenever $B_\varrho \subset B_{2R}$ is a concentric to B_{2R} for the function w in (3.23):

$$\int_{B_\varrho} |w - (w)_{B_\varrho}| dx \leq c \left(\frac{\varrho}{R} \right)^\beta \int_{B_{2R}} |w - (w)_{B_{2R}}| dx + cRs.$$

At this stage combining the previous estimate with (3.25) in the same way as we combined (4.40) with (4.11) in the proof of Theorem 1.4, and iterating as in the proof of Theorem 1.1, we finally come up with (4.43). We just mention that, with the notation introduced in the proof of Theorem 1.1, here we have to define

$$k_i := |(u)_{B_i}|, \quad \text{and} \quad A_i := \int_{B_i} |u - (u)_{B_i}| dx,$$

while no assumption other than measurability is to be assumed on the partial map $x \mapsto a(x, \cdot)$.

5. GENERAL WEAK, AND VERY WEAK, SOLUTIONS

5.1. General elliptic problems. Here we consider Dirichlet problems of the type (1.7) assuming that $\mu \in \mathcal{M}(\Omega)$ is a general Radon measure with finite total mass, and prove that estimate (1.6) holds for general weak and very weak solutions. We are hereby considering the case of zero boundary datum for simplicity, but more general cases can be considered; see also Theorem 5.2 below for general weak (energy) solutions. We need some preliminary terminology. As mentioned in Section 2 uniqueness of general very weak solutions fails, therefore one is led to consider very weak solutions enjoying additional properties, which in some situations are unique; one of such classes, of interest here, is the one of **Solutions Obtained by Limit of Approximations** (SOLA). For the sake of completeness we here recall the approximation procedure in question; the main reference here are the works of Boccardo & Gall  uet [10, 11] and Dall'Aglio [15]. We consider a standard, symmetric and non-negative mollifier $\phi \in C_0^\infty(B_1)$ such that $\|\phi\|_{L^1(\mathbb{R}^n)} = 1$, and then define, for every positive integer h , the mollifier $\phi_h(x) := h^n \phi(hx)$. Finally the functions $\mu_h : \mathbb{R}^n \rightarrow \mathbb{R}$ are defined via convolution, $\mu_h(x) := (\mu * \phi_h)(x)$. Next, by standard monotonicity methods, we find a unique solution $u_h \in W_0^{1,p}(\Omega)$ to

$$(5.1) \quad \begin{cases} -\operatorname{div} a(x, Du_h) = \mu_h & \text{in } \Omega \\ u_h = 0 & \text{on } \partial\Omega. \end{cases}$$

Up to passing to a not relabeled subsequence we may assume that $\mu_h \rightharpoonup \mu$ weakly in the sense of measures, while the results in [10, 11] imply

$$(5.2) \quad Du_h \rightarrow Du \quad \text{strongly in } L^q(\Omega) \text{ for every } q < \frac{n(p-1)}{n-1}, \text{ and a.e.}$$

so that (1.7) is solved by u in the usual distributional sense, and therefore u is a SOLA to (1.7); this of being a limit solution of more regular solutions indeed defines SOLA. Moreover, by [7, 15], in the case $\mu \in L^1(\Omega)$ we also have that u is the only SOLA of (1.7), in the sense that if $v \in W_0^{1,p-1}(\Omega)$ is a distributional solution to (1.7)₁ obtainable as a pointwise limit of solutions $v_h \in W_0^{1,p}(\Omega)$ to problems of the type (5.1) with μ_h replaced by $\tilde{\mu}_h$ and $\tilde{\mu}_h \rightharpoonup \mu$ weakly in $L^1(\Omega)$, then we have $u \equiv v$; for more information on uniqueness see Remark 5.1 below. Now we turn back to the proof of Theorem 1.1, and in particular to estimate (3.55), that we

write when applied to u_h , that is

$$(5.3) \quad k_{m+1}^{(h)} \leq c \left(A_0^{(h)} + k_0^{(h)} + \sum_{i=0}^{m-1} \left[\frac{|\mu_h|(B_i)}{R_i^{n-1}} \right]^{\frac{1}{p-1}} \right) + c \sum_{i=0}^{m-1} [L_1 \omega(R_i)]^{\frac{2}{p}} (k_i^{(h)} + s),$$

where here we have obviously set

$$k_{m+1}^{(h)} := |(Du_h)_{B_{m+1}}| \quad \text{and} \quad A_0^{(h)} := \int_{B_R} |Du_h - (Du_h)_{B_R}| dx$$

compare with (3.50) and (3.52), respectively. We note that the application of estimate (3.55) to u_h is legal since the only thing we used in the proof of (3.55) was that $u \in W_{\text{loc}}^{1,p}(\Omega)$; anyway in the present setting we have $Du \in C^0(\Omega)$ by standard regularity theory. Letting $h \rightarrow \infty$ in (5.3), and using (5.2) and the weak convergence of measures, yields

$$(5.4) \quad k_{m+1} \leq c \left(A_0 + k_0 + \sum_{i=0}^{m-1} \left[\frac{|\mu|(\overline{B_i})}{R_i^{n-1}} \right]^{\frac{1}{p-1}} \right) + c \sum_{i=0}^{m-1} [L_1 \omega(R_i)]^{\frac{2}{p}} (k_i + s),$$

where we are adopting the notation established in the proof of Theorem 1.1. From (5.4) on the rest of the proof follows as for Theorem 1.1 after (3.55), but taking into account essentially two facts: first, the convergence in (3.66) just takes place at Lebesgue points, and therefore almost everywhere; second a slight - but obvious - adjustment (enlargement of the balls) has to be made in order to overcome the presence of $|\mu|(\overline{B_i})$ rather than $|\mu|(B_i)$, and to recover the definition of Wolff potential. We have therefore proved the following:

Theorem 5.1. *Let $u \in W_0^{1,p-1}(\Omega)$ be a SOLA to the problem (1.7) - which is unique in the case $\mu \in L^1(\Omega)$ - under the assumptions (1.2) and (1.3). Then there exists a non-negative constant $c \equiv c(n, p, \nu, L)$, and a positive radius $\tilde{R} \equiv \tilde{R}(n, p, \nu, L, L_1, \omega(\cdot))$ such that estimate (1.6) holds whenever $B(x_0, 2R) \subseteq \Omega$ and $R \leq \tilde{R}$, for almost every $x_0 \in \Omega$. Moreover, when the vector field $a(\cdot)$ is independent of x , estimate (1.6) holds without any restriction on R .*

Remark 5.1. As proved in [7, 57], in the case $p = 2$ the Solutions Obtained by Limit of Approximations (SOLA) of previous theorem are unique in the sense that if $v \in W_0^{1,p-1}(\Omega)$ is a distributional solution to (1.7)₁ obtainable as a pointwise limit of solutions $v_h \in W_0^{1,p-1}(\Omega)$ to problems of the type (5.1) with μ_h replaced by $\tilde{\mu}_h$ and $\tilde{\mu}_h \rightharpoonup \mu$, then we have $u \equiv v$. Moreover, as remarked a few lines above, the same uniqueness result holds when $\mu \in L^1(\Omega)$.

In the case $\mu \in W^{-1,p'}(\Omega)$, we can deal with standard weak solutions, i.e. $u \in W_{\text{loc}}^{1,p}(\Omega)$, by mean of local approximations of the type

$$(5.5) \quad \begin{cases} -\operatorname{div} a(x, Du_h) = \mu_h & \text{in } \Omega' \\ u_h = u & \text{on } \partial\Omega' \end{cases}$$

whenever $\Omega' \Subset \Omega$ are smooth sub-domains, just along the lines of the proof above. In this case the uniqueness of solutions obtainable in the limit is obviously guaranteed as all solutions are of class $W^{1,p}(\Omega')$. We recall that a characterization of those measures μ such that $\mu \in W^{-1,p'}(\Omega)$ is that

$$(5.6) \quad \int_{\Omega} \mathbf{W}_{1,p}^{\mu}(x, 1) d|\mu|(x) = \int_{\Omega} \int_0^1 \left(\frac{|\mu|(B(x, \varrho))}{\varrho^{n-p}} \right)^{\frac{1}{p-1}} \frac{d\varrho}{\varrho} d|\mu|(x) < \infty$$

as proved in [28]; see also [60, Theorem 4.7.5]. In particular a measure $\mu \in \mathcal{M}(\Omega)$ satisfying the density condition

$$|\mu|(B_R) \lesssim R^{n-p+\varepsilon}$$

for some $\varepsilon > 0$ satisfies (5.6); for this case see also [32, 33, 40]. We again recall that here we have trivially extended μ to the whole \mathbb{R}^n . Moreover the approximation scheme in (5.5) is not necessary when μ is an integrable function. We recall that the condition $\mu \in L^{\frac{np}{np-n+p}}(\Omega)$ for $p \leq n$ implies $\mu \in W^{-1,p'}(\Omega)$ simply by Sobolev embedding theorem. In this case the proof of Theorem 1.1 in Section 3 works directly and leads to (1.6) which thereby holds almost everywhere. Summarizing

Theorem 5.2. *Let $u \in W_{\text{loc}}^{1,p}(\Omega)$ be a local weak solution to (1.1), under the assumptions (1.2) and (1.3), and with $\mu \in W^{-1,p'}(\Omega)$. Then there exists a constant $c \equiv c(n, p, \nu, L) > 0$, and a positive radius $\tilde{R} \equiv \tilde{R}(n, p, \nu, L, L_1, \omega(\cdot))$ such that estimate (1.6) holds whenever $B(x_0, 2R) \subseteq \Omega$ and $R \leq \tilde{R}$, for almost every $x_0 \in \Omega$. Moreover, when the vector field $a(\cdot)$ is independent of x , estimate (1.6) holds without any restriction on R .*

Remark 5.2. By recent higher differentiability result for solutions to measure data problems - see [48, Theorem 1.5] - in the last two theorems estimate (1.6) does not only hold for a.e. $x_0 \in \Omega$, but indeed outside a set with Hausdorff measure less than $n - 1$. In the case $\mu \in L_{\text{loc}}^\gamma(\Omega)$, with $1 < \gamma \leq np/(np - n + p)$, the Hausdorff dimension decreases up to $n - \gamma$.

5.2. General parabolic problems. Here, along the lines of the previous section, we give the extension of estimates (1.23) to parabolic SOLA of (1.26) for general measures $\mu \in \mathcal{M}(\Omega_T)$, following this time the approximation schemes in [9, 15]. To this aim we shall consider Cauchy-Dirichlet problems of the type

$$(5.7) \quad \begin{cases} \partial_t u_h - \operatorname{div} a(x, t, Du_h) = \mu_h & \text{in } \Omega_T \\ u_h = 0 & \text{on } \partial_{\text{par}} \Omega_T, \end{cases}$$

solved by a functions $u_h \in C^0(-T, 0; L^2(\Omega)) \cap L^2(-T, 0; W_0^{1,2}(\Omega))$, $h \in \mathbb{N}$. As in the elliptic case μ_h is a sequence of smooth functions obtained from μ via mollification, and therefore such that $\mu_h \rightharpoonup \mu$ weakly in the sense of measures. The main convergence result in [9] claims that, up to extracting a not-relabeled subsequence, we have that there exists a function $u \in L^r(-T, 0; W^{1,q}(\Omega))$ such that $u_h \rightharpoonup u$ weakly in $L^r(-T, 0; W^{1,q}(\Omega))$, for every choice of the parameters

$$(5.8) \quad q < \frac{n}{n-1}, \quad r \in [1, 2], \quad \frac{2}{r} + \frac{n}{q} > n + 1,$$

see the assumptions of [9, Theorem 1.2], while the result is proved in [9, Section 3]. In particular we can choose $q = r > 1$ still matching (5.8) and therefore the main convergence result in [9, Theorem 3.3] states that, always up to a not-relabeled subsequence, $Du_h \rightharpoonup Du$ a.e. in Ω_T . Then, since the sequence $\{Du_h\}$ is still bounded in $L^q(\Omega_T)$ we infer that $Du_h \rightharpoonup Du$ strongly in $L^1(\Omega_T)$, and finally u weakly solves (1.26) in the sense of (2.1); in other words u is a SOLA to (1.26). We recall that uniqueness of SOLA still holds in the parabolic case provided $\mu \in L^1(\Omega_T)$; for this we refer to [15]. At this stage, we can repeat the strategy of the elliptic proof of Theorem 5.1: i.e. we go to the proof of Theorem 1.2, we apply it to every solutions u_h in order to have inequalities as in (5.3), we pass to the limit with respect to $h \in \mathbb{N}$ the resulting estimates on the numbers $k_m^{(h)} \equiv k_m$ - compare with (5.3) - and then we conclude with the desired pointwise inequality. Summarizing

Theorem 5.3. *Let $u \in L^1(-T, 0; W^{1,1}(\Omega))$ be a SOLA to the problem (1.26) - which is unique in the case $\mu \in L^1(\Omega_T)$ - under the assumptions (1.18) and (1.3). Then there exists a constant $c \equiv c(n, \nu, L)$ and a radius $\tilde{R} \equiv \tilde{R}(n, \nu, L, L_1, \omega(\cdot))$ such that estimate (1.23) holds whenever $Q(x_0, t_0; 2R) \subseteq \Omega$, for almost every $(x_0, t_0) \in$*

Ω_T , provided $R \leq \tilde{R}$. When the vector field $a(\cdot)$ is independent of the space variable x , estimate (1.23) holds without any restriction on R .

We similarly have

Theorem 5.4. *Let $u \in L^1(-T, 0; W^{1,1}(\Omega))$ be a SOLA to the problem (1.26) - which is unique in the case $\mu \in L^1(\Omega_T)$ - under the assumptions (1.28). Then there exists a constant $c \equiv c(n, \nu, L)$ and a radius $\tilde{R} \equiv \tilde{R}(n, \nu, L, L_1, \omega(\cdot))$ such that estimate (1.23) holds whenever $Q(x_0, t_0; 2R) \subseteq \Omega$, for almost every $(x_0, t_0) \in \Omega_T$, provided $R \leq \tilde{R}$. When the vector field $a(\cdot)$ is independent of the space variable x , estimate (1.23) holds without any restriction on R .*

Moreover, when μ is regular enough - for instance when it belongs to the dual space $L^2(-T, 0; W^{-1,2}(\Omega))$ - estimate (1.23) works for general weak solutions $u \in L^2(-T, 0; W^{1,2}(\Omega))$ to (1.17).

6. INTEGRABILITY ESTIMATES

In this section we rapidly show how to get the integrability estimates directly from pointwise estimates as (1.6) and (1.23); the most interesting results are probably those in the parabolic case, since they are completely new, as described below, while in the elliptic case we shall mostly find a unified approach to the proof of several results appearing in the literature. In both the elliptic and the parabolic case we shall confine ourselves to outline the strategy of proof, since then the details can be easily added by the interested reader.

A main ingredient in the elliptic case here is of course the possibility to control the non-linear Wolff type potentials via the non-linear Havin-Maz'ja potential, and ultimately via Riesz potentials, which is described in (1.15). Therefore the main property to use is the boundedness of the Riesz potential in Lebesgue spaces

$$(6.1) \quad \|I_\beta(g)\|_{L^{\frac{nq}{n-\beta q}}(\mathbb{R}^n)} \leq c \|g\|_{L^q(\mathbb{R}^n)} \quad q \in (1, n/\beta),$$

with similar estimates to hold also in Marcinkiewicz and Orlicz spaces via interpolation; see also [48] and next section. Using (6.1) and (1.15) we then come to various mapping properties as for instance

$$(6.2) \quad \mu \in L^q \implies \mathbf{W}_{\frac{1}{p}, p}^\mu(\cdot, R) \in L^{\frac{nq(p-1)}{n-q}} \quad q \in (1, n).$$

The last implication comes of course with related quantitative estimates. In turn, using (6.2) in combination with (1.6) leads to deduce the following inclusion for solutions to (1.7) considered in Theorems 5.1-5.2:

$$(6.3) \quad \mu \in L^q \implies Du \in L_{\text{loc}}^{\frac{nq(p-1)}{n-q}}(\Omega) \quad q \in (1, n),$$

The last implication allows to recover various estimates available in the literature [10, 11, 18, 29] which are usually achieved via very different techniques, according to the size of q . There is now a large number of possible variants to the last result. In fact by the known bounds on the Riesz potential, it is possible to easily find virtually *all kinds of rearrangement invariant function spaces integral estimates*. For more details on this we refer to the next section, which is dedicated to parabolic problems, and to Remark 6.1 below. Further developments of our viewpoint can be found in the forthcoming paper of Cianchi [14], where a few interesting consequences of estimate (1.6) are presented. For instance, in [14] the following consequence of (1.8) is presented

$$(6.4) \quad \mu \in L\left(n, \frac{1}{p-1}\right) \implies Du \in L^\infty \quad \text{locally in } \Omega.$$

In this case something more can be actually asserted, namely the continuity of Du , as will be proved in the subsequent paper [23]. Again, Wolff potentials based continuity criteria for the gradient of solutions can be obtained starting from the methods presented here, and will be presented in [23].

We also mention that by *localizing* inclusions (6.2) and (6.3) it is also possible to obtain explicit local estimates related to the regularity results mentioned above, something which is also not easily reachable via the global techniques usually used in the literature. For this we also refer to [47, 48, 48].

Remark 6.1 (Borderline cases). In the regularity theory of measure data problems, for a long while it has been an open issue to establish the validity of estimates in Marcinkiewicz spaces in the conformal case $p = n$. Specifically, for a SOLA solution u of (1.1) we have the inclusion $Du \in \mathcal{M}^n(\Omega)$ - see (6.10) below for the definitions. This has been first proved in [19] directly for the p -Laplacean system, see also [47] for different approaches. Here we would like to outline how estimates (1.6)-(1.16) immediately imply that $Du \in \mathcal{M}_{\text{loc}}^n(\Omega, \mathbb{R}^n)$ in our setting. This goes via analyzing properties of Wolff potentials in Lorentz and Marcinkiewicz spaces - see (6.9)-(6.10) below for the definitions; indeed by standard properties of Riesz potentials - see for instance [48, Section 4] - and (1.15) we have

$$\mu \in \mathcal{M}(\Omega) \implies \mathbf{W}_{\frac{1}{p}, p}^\mu(\cdot, R) \in \mathcal{M}^{\frac{n(p-1)}{n-1}}.$$

More in general, as in (6.2) we have that (1.15) implies

$$(6.5) \quad \mu \in L(q, \gamma) \implies \mathbf{W}_{\frac{1}{p}, p}^\mu(\cdot, R) \in L\left(\frac{nq(p-1)}{n-q}, \gamma(p-1)\right)$$

whenever $q \in (1, n)$ and $\gamma \in (0, \infty]$. Another open regularity problem was the borderline L^p -regularity of solutions of SOLA to (1.4), this means establishing, in the case $p < n$, that

$$(6.6) \quad \mu \in L\left(\frac{np}{np-n+p}, \gamma\right) \implies Du \in L(p, \gamma(p-1))$$

whenever $\gamma \in (0, \infty]$. This problem, raised several times in the literature - see for instance [8, 34] and related references - has been recently settled in [48, Theorem 2]. Here another proof of (6.6) follows directly from estimates (1.6)-(1.16) via (6.5).

Remark 6.2 (Sharpness). Inequality (1.6) is in a sense optimal as $\mathbf{W}_{\frac{1}{p}, p}^\mu$ allows to recast all the integral estimates in the scale of Lebesgue and Lorentz spaces as described in the previous remark. This is obvious in the case $p = 2$ due to well-known representation formulas for the Poisson equation (1.12); the optimality for $p > 2$ follows since the result in (6.3) is the best possible; at this point a better potential in (1.6) would imply a better estimate in (6.3), which is indeed impossible. In other words, in (1.6) no potential of the type $\mathbf{W}_{\beta, p}^\mu$ with $\beta > 1/p$ can replace $\mathbf{W}_{\frac{1}{p}, p}^\mu$, for every $p \geq 2$. Let us now test estimate (1.6) in the case we have the following Dirichlet problem for $p \leq n$:

$$(6.7) \quad \begin{cases} -\Delta_p u &= \delta_n & \text{in } B_1 \\ u &= 0 & \text{on } \partial B_1 \end{cases}$$

with $\delta_n := \sigma_{n-1}\delta$, where δ is the Dirac measure charging the origin and $\sigma_{n-1} := \mathcal{H}^{n-1}(\partial B_1)$ is the $(n-1)$ -dimensional Hausdorff measure of ∂B_1 . The unique solution to (6.7) is given by the so called fundamental solution of the p -Laplacean

operator, that is

$$u(x) := \begin{cases} \frac{p-1}{n-p} \left(|x|^{\frac{p-n}{p-1}} - 1 \right) & \text{if } p < n \\ -\log |x| & \text{if } p = n. \end{cases}$$

The uniqueness is a consequence of classical results of Serrin [54] and Kichenassamy & Veron [31]; an account of the proof can be found in [50, Section 4.4]. Let us show that in this case estimate (1.6) reverses: let us take $x_0 \neq 0$ such that $|x_0| \leq 1/2$, then we have

$$(6.8) \quad \int_{B(x_0, R)} |Du| dx + \mathbf{W}_{\frac{1}{p}, p}^{\delta_n}(x_0, 2R) \leq c |Du(x_0)|,$$

with $R = |x_0|$, and for a constant c depending only on n and p . Indeed by co-area formula we have

$$\begin{aligned} \int_{B(x_0, R)} |Du| dx &\leq c(n) \int_{B(0, 2R)} |Du| dx \\ &= \frac{c(n)}{R^n} \int_0^{2R} \int_{\partial B_\varrho} |Du(y)| d\mathcal{H}^{n-1}(y) d\varrho \\ &= \frac{c(n, p)}{R^n} \int_0^{2R} \varrho^{n+\frac{1-n}{p-1}} \frac{d\varrho}{\varrho} \\ &= c(n, p) R^{\frac{1-n}{p-1}} \leq c(n, p) |Du(x_0)|. \end{aligned}$$

On the other hand, as $\delta_n(B(x_0, \varrho)) = 0$ for $\varrho \leq R$ and $\delta_n(B(x_0, \varrho)) = 1$ otherwise, we also have

$$\mathbf{W}_{\frac{1}{p}, p}^{\delta_n}(x_0, 2R) \leq c(n, p) \int_R^{2R} \varrho^{\frac{1-n}{p-1}} \frac{d\varrho}{\varrho} \leq c(n, p) R^{\frac{1-n}{p-1}} \leq c(n, p) |Du(x_0)|,$$

so that (6.8) follows combining the last two inequalities. Inequalities as (6.8) do not only hold when the right hand side of the equation is a Dirac measure; indeed, when considering more general measures concentrating on lower dimensional set – lines, hyper-planes, etc – we observe power-type singularities of solutions and estimates like (6.8).

6.1. Parabolic integral estimates. In the parabolic case the consequences of the pointwise estimates are more interesting as we are able to get sharp borderline estimates in Lorentz and Marcinkiewicz spaces which do not seem immediately reachable via the known parabolic techniques. Moreover, as in the elliptic case, we catch directly borderline cases in Lorentz spaces. To this aim let us recall that a map $g : A \subseteq \mathbb{R}^{n+1} \rightarrow \mathbb{R}^k$, with A being an open subset and $k \in \mathbb{N}$, and for $q \in [1, \infty)$ and $\gamma \in (0, \infty]$, is said to belong to the Lorentz space $L(q, \gamma)(A)$ iff:

$$(6.9) \quad \|g\|_{L(q, \gamma)(A)}^\gamma := \gamma \int_0^\infty (\lambda^q |\{x \in A : |g(x)| > \lambda\}|)^{\frac{\gamma}{q}} \frac{d\lambda}{\lambda} < \infty \quad \text{for } \gamma < \infty.$$

In the case $\gamma = \infty$ we have instead Marcinkiewicz spaces $\mathcal{M}^q(A) \equiv L(q, \infty)(A)$:

$$(6.10) \quad \sup_{\lambda > 0} \lambda^q |\{x \in A : |g(x)| > \lambda\}| =: \|g\|_{\mathcal{M}^q(A)}^q < \infty.$$

The local variants of such spaces are then defined in the obvious way. We here just recall that Lorentz spaces “tune” Lebesgue spaces via the second index according to the following chain of inclusions:

$$L^r \equiv L(r, r) \subset L(q, \gamma) \subset L(q, q) \subset L(q, r) \subset L(\gamma, \gamma) \equiv L^\gamma,$$

which hold whenever $0 < \gamma < q < r \leq \infty$. The following inequalities can now be obtained using the definition of caloric Riesz potentials given in (1.20):

$$(6.11) \quad \|I_\beta(g)\|_{L(\frac{Nq}{N-\beta q}, \gamma)(\mathbb{R}^{n+1})} \leq c\|g\|_{L(q, \gamma)(\mathbb{R}^{n+1})}$$

which holds whenever $q > 1$ and $\beta q < N$, and, in the case $q = 1$,

$$(6.12) \quad \|I_\beta(g)\|_{\mathcal{M}^{\frac{N}{N-\beta}}(\mathbb{R}^{n+1})} \leq c|g|(\mathbb{R}^{n+1}).$$

The last inequality still holds when g is a measure and $\beta < N$. Inequalities (6.11) and (6.12) can be localized, more precisely it is not difficult to see that $I_\beta(\cdot)$ controls the localized potential $\mathbf{I}_1^\mu(\cdot; R)$ defined in (1.22) in the sense that $\mathbf{I}_\beta^\mu(\cdot; R) \lesssim I_\beta(g)$, holds. Therefore we have that

$$(6.13) \quad \mathbf{I}_\beta^\mu(\cdot; R): L(q, \gamma) \rightarrow L\left(\frac{Nq}{N-\beta q}, \gamma\right) \quad \text{and} \quad \mathbf{I}_\beta^\mu(\cdot; R): L^1 \rightarrow \mathcal{M}^{\frac{N}{N-\beta}}$$

hold, continuously, assuming $\beta q < N$ and $\beta < N$, respectively. The estimates in (6.13) refer to \mathbb{R}^{n+1} as supporting domain, but they can be localized using cut-off functions and therefore yield local estimates for caloric Riesz potentials; see for instance [48, 49]. The final outcome is the following result:

Theorem 6.1. *Let $u \in L^1(-T, 0; W^{1,1}(\Omega))$ be a SOLA to the problem (1.26) - which is unique in the case $\mu \in L^1(\Omega_T)$ - under the assumptions (1.18) and (1.3). Then the following implications hold:*

$$(6.14) \quad \mu \in L(q, \gamma) \implies Du \in L\left(\frac{Nq}{N-q}, \gamma\right) \quad \text{for } 1 < q < N \text{ and } 0 < \gamma < \infty$$

and

$$(6.15) \quad \mu \in \mathcal{M}(\Omega) \implies Du \in \mathcal{M}^{\frac{N}{N-1}}$$

with all the previous inclusions being meant as **local** in Ω .

The previous result is of course a straightforward consequence of estimate (1.23) together with the properties in (6.13). Moreover, using a localized version of the estimates relative to (6.13) it is also possible to obtain *explicit local estimates* relative to the implications in (6.14)-(6.15). See for instance the strategies adopted in [48].

REFERENCES

- [1] Acerbi E. & Mingione G.: Regularity results for a class of functionals with non-standard growth. *Arch. Ration. Mech. Anal.* 156 (2001), 121–140.
- [2] Acerbi E. & Mingione G.: Gradient estimates for a class of parabolic systems. *Duke Math. J.* 136 (2007), 285–320.
- [3] Adams D. R. & Bagby R. J.: Translation-dilation invariant estimates for Riesz potentials. *Indiana Univ. Math. J.* 23 (1974), 1051–1067.
- [4] Adams D. R. & Hedberg L. I.: *Function spaces and potential theory*. Grundlehren der Mathematischen Wissenschaften 314. Springer-Verlag, Berlin, 1996.
- [5] Adams D. R. & Meyers N. G.: Thinness and Wiener criteria for non-linear potentials. *Indiana Univ. Math. J.* 22 (1972), 169–197.
- [6] Benilan P. & Boccardo L. & Gallouët T. & Gariepy R. & Pierre M. & Vázquez J. L.: An L^1 -theory of existence and uniqueness of solutions of nonlinear elliptic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (IV)* 22 (1995), 241–273.
- [7] Boccardo L.: Problemi differenziali ellittici e parabolici con dati misure [Elliptic and parabolic differential problems with measure data]. *Boll. Un. Mat. Ital. A (7)* 11 (1997), 439–461.
- [8] Boccardo L.: Marcinkiewicz estimates for solutions of some elliptic problems with nonregular data. *Ann. Mat. Pura Appl. (V)* 188 (2009), 591–601.
- [9] Boccardo L. & Dall’Aglio A. & Gallouët T. & Orsina L.: Nonlinear parabolic equations with measure data. *J. Funct. Anal.*, 147 (1997), 237–258.
- [10] Boccardo L. & Gallouët T.: Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.* 87 (1989), 149–169.

- [11] Boccardo L. & Gallouët T.: Nonlinear elliptic equations with right-hand side measures. *Comm. Partial Differential Equations* 17 (1992), 641–655.
- [12] Bögelein V. & Haberman J.: Gradient estimates via non-standard potentials. *Preprint 2009*.
- [13] Cianchi A.: Maximizing the L^∞ norm of the gradient of solutions to the Poisson equation. *J. Geom. Anal.* 2 (1992), 499–515.
- [14] Cianchi A.: Non-linear potentials, local solutions to elliptic equations and rearrangements. *In preparation*.
- [15] Dall'Aglio A.: Approximated solutions of equations with L^1 data. Application to the H -convergence of quasi-linear parabolic equations. *Ann. Mat. Pura Appl. (IV)* 170 (1996), 207–240.
- [16] DiBenedetto E.: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Anal. TMA* 7 (1983), 827–850.
- [17] DiBenedetto E.: *Degenerate parabolic equations*. Universitext. Springer-Verlag, New York, 1993.
- [18] DiBenedetto E. & Manfredi J.J.: On the higher integrability of the gradient of weak solutions of certain degenerate elliptic systems. *Amer. J. Math.* 115 (1993), 1107–1134.
- [19] Dolzmann G. & Hungerbühler N. & Müller S.: Uniqueness and maximal regularity for non-linear elliptic systems of n -Laplace type with measure valued right hand side. *J. Reine Angew. Math.* 520 (2000), 1–35.
- [20] Duzaar F. & Mingione G.: The p -harmonic approximation and the regularity of p -harmonic maps. *Calc. Var. Part. Diff. Equ.* 20 (2004), 235–256.
- [21] Duzaar F. & Mingione G.: Regularity for degenerate elliptic problems via p -harmonic approximation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 21 (2004), 735–766.
- [22] Duzaar F. & Mingione G.: Gradient estimates in non-linear potential theory. *Rend. Lincei - Mat. Appl.* 20 (2009), 179–190.
- [23] Duzaar F. & Mingione G.: Gradient continuity estimates. *Submitted 2009*.
- [24] Giusti E.: *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
- [25] Grüter M. & Widman K.O.: The Green function for uniformly elliptic equations. *manuscripta math.* 37 (1982), 303–342.
- [26] Hamburger C.: Regularity of differential forms minimizing degenerate elliptic functionals. *J. Reine Angew. Math. (Crelles J.)* 431 (1992), 7–64.
- [27] Havin M. & Maz'ja V. G.: A nonlinear potential theory. *Russ. Math. Surveys* 27 (1972), 71–148.
- [28] Hedberg L.I. & Wolff T.: Thin sets in nonlinear potential theory. *Ann. Inst. Fourier (Grenoble)* 33 (1983), 161–187.
- [29] Iwaniec T.: Projections onto gradient fields and L^p -estimates for degenerated elliptic operators. *Studia Math.* 75 (1983), 293–312.
- [30] Jin T. & Mazya V. & Van Schaftingen J.: Pathological solutions to elliptic problems in divergence form with continuous coefficients. *Comptes Rendus Mathématique* 347 (2009), 773–778.
- [31] Kichenassamy S. & Verón L.: Singular solutions to the p -Laplace equation. *Math. Ann.* 275 (1986), 599–615.
- [32] Kilpeläinen T.: Hölder continuity of solutions to quasilinear elliptic equations involving measures. *Potential Anal.* 3 (1994), 265–272.
- [33] Kilpeläinen T.: Singular solutions to p -Laplacian type equations. *Ark. Mat.* 37 (1999), 275–289.
- [34] Kilpeläinen T. & Li G.: Estimates for p -Poisson equations. *Diff. Int. Equ.* 13 (2000), 791–800.
- [35] Kilpeläinen T. & Malý J.: The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.* 172 (1994), 137–161.
- [36] Kristensen J. & Mingione G.: The singular set of minima of integral functionals. *Arch. Ration. Mech. Anal.* 180 (2006), 331–398.
- [37] Kuusi T. & Mingione G.: Non-linear Nash density estimates. *In preparation*.
- [38] Labutin D.: Potential estimates for a class of fully nonlinear elliptic equations. *Duke Math. J.* 111 (2002), 1–49.
- [39] Lieberman G.M.: The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations. *Comm. Part. Diff. Equ.* 16 (1991), 311–361.
- [40] Lieberman G. M.: Sharp forms of estimates for subsolutions and supersolutions of quasilinear elliptic equations involving measures. *Comm. Part. Diff. Equ.* 18 (1993), 1191–1212.
- [41] Lieberman G. M.: *Second order parabolic differential equations*. World Scientific press, River Edge, 1996.
- [42] Lindqvist P.: On the definition and properties of p -superharmonic functions. *J. reine angew. Math. (Crelles J.)* 365 (1986), 67–79.

- [43] Lindqvist P.: Notes on the p -Laplace equation. *Univ. Jyväskylä, Report* 102, (2006).
- [44] Lindqvist P. & Manfredi J.J.: Note on a remarkable superposition for a nonlinear equation. *Proc. AMS* 136 (2008), 133–140.
- [45] Manfredi J.J.: Regularity for minima of functionals with p -growth. *J. Differential Equations* 76 (1988), 203–212.
- [46] Manfredi J.J.: Regularity of the gradient for a class of nonlinear possibly degenerate elliptic equations. *Ph.D. Thesis*. University of Washington, St. Louis.
- [47] Mingione G.: The Calderón-Zygmund theory for elliptic problems with measure data. *Ann. Scu. Norm. Sup. Pisa Cl. Sci. (5)* 6 (2007), 195–261.
- [48] Mingione G.: Gradient estimates below the duality exponent. *Mathematische Annalen* DOI: 10.1007/s00208-009-0411-z
- [49] Mingione G.: Gradient potential estimates. *J. Europ. Math. Soc.*, to appear.
- [50] Mingione G.: Towards a non-linear Calderón-Zygmund theory. *Quaderni di Matematica*, to appear.
- [51] Moser J.: A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.* 17 1964 101–134. Corrections in: *Comm. Pure Appl. Math.* 20 (1967), 231–236.
- [52] Nash J.: Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* 80 (1958), 931–954.
- [53] Naumann J. & Wolf J.: Interior integral estimates on weak solutions of nonlinear parabolic systems. *Preprint 94-11 Inst. für Math., Humboldt Universität, Bonn* (1994).
- [54] Serrin J.: Local behavior of solutions of quasi-linear equations. *Acta Math.* 111 (1964), 247–302.
- [55] Talenti G.: Elliptic equations and rearrangements. *Ann. Scu. Norm. Sup. Pisa Cl. Sci. (4)* 3 (1976), 697–717.
- [56] Trudinger N.S. & Wang X.J.: On the weak continuity of elliptic operators and applications to potential theory. *Amer. J. Math.* 124 (2002), 369–410.
- [57] Trudinger N.S. & Wang X.J.: Quasilinear elliptic equations with signed measure data. *Disc. Cont. Dyn. Systems A* 23 (2009), 477–494.
- [58] Uhlenbeck K.: Regularity for a class of non-linear elliptic systems. *Acta Math.* 138 (1977), 219–240.
- [59] Ural'tseva N.N.: Degenerate quasilinear elliptic systems. *Zap. Na. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 7 (1968), 184–222.
- [60] Ziemer W. P.: *Weakly differentiable functions. Sobolev spaces and functions of bounded variation*. Graduate Texts in Mathematics, 120. Springer-Verlag, New York, 1989.

FRANK DUZAAR, DEPARTMENT MATHEMATIK, UNIVERSITÄT ERLANGEN-NÜRNBERG, BISMARCKSTRASSE 1 1/2, 91054 ERLANGEN, GERMANY
E-mail address: duzaar@mi.uni-erlangen.de

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PARMA, VIALE G. P. USBERTI 53/A, CAMPUS, 43100 PARMA, ITALY
E-mail address: giuseppe.mingione@unipr.it.